Investigating the role of ionic networks in the thermostability of citrate synthase from the Archaea

Michael, Rebecca Anne

Award date:
2002

Awarding institution:
University of Bath

INVESTIGATING THE ROLE OF IONIC NETWORKS IN THE THERMOSTABILITY OF CITRATE SYNTHASE FROM THE ARCHAEA

Submitted by Rebecca Anne Michael

for the degree of PhD of the University of Bath

2002

Copyright

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

DEDICATED

TO

MAM AND DAD

ABSTRACT

Pyrococcus furiosus (*Pf*) is an hyperthermophilic Archaeon that has a growth optimum of 100°C. The crystal structure of its citrate synthase (CS) has been determined and, compared to less thermophilic and mesophilic CS structures, it is at the subunit interface of the dimeric enzyme that interesting changes in the structure are observed. These changes are thought to be structural features of protein thermostability and, as citrate synthase requires dimer integrity for catalytic activity, subunit interactions that maintain the integrity of the dimer are important. In the *Pf*CS, the C-terminal region of each monomer folds over the surface of the other monomer, providing a type of subunit interaction that is unique to this enzyme. There is also an increase in the number of inter-subunit ion pairs in the *Pf*CS compared with less thermostable CSs. A major interaction at the subunit interface of the *Pf*CS is a 5-residue ion-pair network involving amino acids from helices G and M. *Pf*CS mutants were created, by site-directed mutagenesis, to disrupt these subunit interactions by the complete and partial removal of the C-terminal arm regions, and substitution of Asp113 to disrupt the ionic network. Determination of the kinetic parameters of these mutants revealed catalytic activities comparable to that of the *Pf*CS wild-type (wt), although thermal inactivation and temperature optima studies show that the *Pf*CS mutants are less thermostable than the *Pf*CS wt.

Thermoplasma acidophilum (*Ta*) is a thermophilic Archaeon with a growth optimum of 55°C. Compared to the crystal structure of *Pf*CS, *Ta*CS does not contain a 5-residue ionic network at the subunit interface. Attempts were made to construct the *Pf*CS ionic network into the *Ta*CS by site-directed mutagenesis. All mutations were made in

different combinations. Although all of the *Ta*CS mutants had catalytic activities comparable to that of the *Ta*CS wt, thermal inactivation and temperature optima studies revealed that all mutants, apart from the D94H single mutant, were less thermostable than the *Ta*CS wt. Preliminary study of the *Ta*CS D94H mutant has suggested that this mutation has created an extra inter-subunit ionic bond equivalent to an ionic bond found in the citrate synthase of *Sulfolobus solfataricus* (growth optima of 87°C).

CONTENTS

CHAPTER 2

Materials and Methods

CHAPTER 3

Investigating the role of the inter-subunit ionic network and the C-terminal region in the thermal stability of *Pyrococcus furiosus* citrate synthase.

CHAPTER 4

Construction of inter-subunit ionic network mutants of *Thermoplasma acidophilum* citrate synthase

CHAPTER 5

Kinetic and thermal characterisation of the *Thermoplasma acidophilum* citrate synthase ionic network mutants

ACKNOWLEDGEMENTS

I would like to thank my supervisors Professor Mike Danson and Dr David Hough for their assistance and guidance throughout my PhD, and for all the useful and interesting discussions. A big thank-you is also due for their invaluable input in the production of this manuscript.

I would also like to thank everyone in the Centre For Extremophile Research for all their support and making work a fun place to be. Special thanks to Dr Frank Sugar for his support throughout the three years and for being the best house-mate ever, and to Dr Carl Thompson for being my constant source of practical advice. I would also like to give a big thank-you to Dr Susan Crennell for the production of all of the help with the figures in this thesis (and apologise for all the pestering!).

I gratefully acknowledge the BBSRC who provided the sponsorship for this project.

A big thank-you to all my family (I can finally get a proper job!), especially to my mum and dad for all their financial and emotion support throughout my time at University. I would also like to thank Matt, Simon, Gary, Claire, and James 'Randy' Randerson for keeping me sane for three years. Finally I would like to thank James for his constant encouragement and support for the past three years.

ABBREVIATIONS

Å	Angström
AcCoA	Acetyl coenzyme A
Asn	Asparagine
Asp	Aspartate
bp	Base pairs
BSA	Bovine serum albumin
CBADH	*Clostridium beijerinckii* alcohol dehydrogenase
CoA	Coenzyme A
CS	Citrate synthase
CyP	Cyclophilin
Cys	Cysteine
ΔG	Gibbs free-energy difference
ΔH	enthalpic contribution to ΔG
TΔS	entropic contribution to ΔG
DNA	Deoxyribonucleic acid
DSC	Differential scanning calorimetry
dm	Double mutant
DTNB	5, 5'-di-thio-2-nitrobenzoic acid
EDTA	Ethylenediamine tetra-acetate
EPPS	N-[2-hydroxyethyl] piperazine-N'-[3-propanesulfonic acid]
F	Folded
GDH	Glutamate dehydrogenase
G/D/I/P	Gly112/Asp113/Ile114/Pro115 four-residue loop in *Pf*CS

Gln	Glutamate
His	Histidine
Ile	Isoleucine
IPMDH	3-isopropylmalate dehydrogenase
kb	Kilobase pairs
kcal	Kilocalorie
kDa	Kilodalton
LB	Luria Bertani medium
Lys	Lysine
Met	Methionine
Mr	Relative molecular mass
$NAD(P)^+$	Nicotinamide phosphate
dNTP	Deoxynucleoside triphosphate
OAA	Oxaloacetate
ORPTase	ororate phosphoribosyl tranferase
PCR	Polymerase chain reaction
Pf	*Pyrococcus furiosus*
PPIase	Peptidyl proline *cis/trans* isomerase
Pro	Proline
R	Gas constant (8.314J/mol/K)
SAP	Shrimp alkaline phosphate
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylmide gel electrophoresis
sm	Single mutation

Ss	*Sulfolobus solfataricus*
T	absolute temperature (K)
TAE	40mM Tris-acetate, 1mM EDTA
TBADH	*Thermoanaerobacter brockii* alcohol dehydrogenase
Taq	*Thermus aquaticus*
Ta	*Thermoplasma acidophilum*
TEMED	Tetramethylethylenediamine
tm	triple mutant
T_{opt}	Temperature optimum
U	Unfolded
wt	Wild-type

CHAPTER 1

Introduction

1.1 Extremophilic organisms

Extremophiles are organisms that have evolved to exist in a variety of extreme environments and they fall into a number of different classes that include thermophiles (optimal growth at 55-80°C), hyperthermophiles (80-110°C), acidophiles (pH 0-4), alkalophiles (optimal growth at pH 9-12), psychrophiles (–2-10°C), halophiles (2-5M NaCl), piezophiles (growth in >400 atmospheres), and radiation resistant and heavy metal tolerant organisms.

The most extremely tolerant of these microrganisms belong to a distinct domain of life called the Archaea, identified in 1977 by Carl Woese, although many extremophilic bacteria have also been described. The Archaea represent a third line of living organisms in addition to Eukarya and Bacteria [Woese *et al.*, 1990]. Archaea exhibit a mixture of bacterial, eukaryotic and unique features at the molecular level [reviewed by Forterre, 1997]. Interest in the Archaea has primarily been stimulated by the ability of many to occupy extreme environmental niches, which may provide clues about early evolutionary divergence [reviewed by Keeling & Doolittle, 1995]. The extreme conditions in which the Archaea are found impose phenotypes apparently well suited to the type of environment thought to exist during early life on earth.

1.2 Extremozymes and their industrial applications

The discovery of new extremophilic micro-organisms and their enzymes, extremozymes, has had a great impact on the field of biotechnology as mesophilic enzymes are often not suited to the harsh conditions required in industrial processes. Hence, extremophiles are often routinely used as the source of novel stable enzymes. Apart from hyperthermophilic extremozymes, which are almost exclusively archaeal in origin, extremozymes that currently find applications originate from both bacterial and archaeal sources, and of these, the majority come from extremophilic bacteria. This bias is related to the fact that current understanding and molecular tools for working with Bacteria are better developed. However, with an ever increasing understanding of Archaea and with ongoing development of improved techniques for studying these organisms, this situation may soon change [discussed by Eichler, 2001]. Recently there has been an increase in the number of genomes from extremophilic organisms that have been sequenced, including the thermophilic Archaea *Thermoplasma acidophilum* [Ruepp *et al.*, 2000] and *Sulfolobus solfataricus* [She *et al.*, 2001], and the hyperthermophilic Archaeon *Pyrococcus furiosus* [Robb *et al.*, 2001]. The availability of these, and many other, genome sequences will aid the search for novel extremozymes of commercial interest.

Recent advances in molecular biology techniques and protein engineering have enabled the production or improvement of extremophilic enzymes by rational design/site-directed mutagenesis and directed evolution have now become common tools for introducing and

improving extremophilic characteristics in both mesophilic and extremophilic enzymes. These techniques are described and discussed in detail in Section 1.6.

Most work has been devoted to thermophilic and hyperthermophilic enzymes, discussed in Section 1.3, although many other extremozymes have important biotechnological potential e.g. psychrophilic and halophilic enzymes. Structural analysis of a cold-active citrate synthase enzyme revealed that catalytic efficiency appears to be achieved by a more accessible active site and an increase in the relative flexibility compared to the hyperthermophilic homologue [Russell *et al.*, 1998]. Sequence analysis of 21 psychrophilic enzymes from different families revealed similar features of cold-adaptation [Gianese *et al.*, 2001], and it is believed that this increase in flexibility allows the enzymes to be catalytically efficient at lower temperatures [reviewed by Lonhienne *et al.*, 2000]. These features can be utilised for biotechnological processes [reviewed by Gerday *et al.*, 2000], for example, in the detergent and food industries, and for the production of heat-labile products.

Study of the structural and biochemical properties of halophilic enzymes has shown that the surfaces of these proteins have a high negative charge which makes them more soluble and renders them more flexible at high salt concentrations [reviewed by Mevarech *et al.*, 2000]. The high surface charge is neutralised by the binding of hydrated salt ions. It has also been found that the high salt concentration requirement of halophilic enzymes is due to a low affinity binding of the salt to specific sites on the surface of the folded protein, thus stabilising the active conformation of the protein [reviewed by Madern *et al.*, 2000]. Not

only are halophilic enzymes extremely salt-tolerant, in general they are able to retain stability at ambient temperatures for prolonged periods and are often thermotolerant. This gives a potential biotechnological advantage of these enzymes over their mesophilic counterparts. Furthermore, due to the nature of the protein surface, halophilic enzymes can potentially retain catalytic activity in environments of low water activity e.g. organic solvents. Such features of these enzymes are of potential use in the production of fermented food and food supplements and in the degradation or transformation of organic pollutants [reviewed by Margesin *et al.*, 2001].

As previously mentioned, heat-tolerant enzymes are currently the most investigated of all extremozymes, not only due to their significant biotechnological potential, but because they also serve as excellent models for the study of protein stability [detailed reviews by Bruins *et al.*, 2001; Sterner & Leibl, 2001; Vieille & Zeikus, 1996; 2001]. This chapter describes the basis of protein thermostability and discusses the industrial and academic interest in thermophilic and hyperthermophilic enzymes.

1.3 Thermostable enzymes and their biotechnological potential

The main reason for selecting enzymes from thermophiles for use in industrial processes is, apart from their high thermal stability, thermostable enzymes also show increased stability under other extreme conditions i.e. high pH or low water concentrations. Stable enzymes are also more likely to allow the use of organic solvents and detergents and are more resistant to proteolytic attack. The main advantages of performing processes at higher

temperatures are reduced risk of microbial contamination, lower viscosity, improved mass transfer rates, and improved solubility of substrates [reviewed by Bruins *et al.*, 2001; Eichler, 2001]. Hence, increasing the temperature of a biotechnological process and using enzymes that can withstand these elevated temperatures, and sometimes remain active depending on the requirements of the industrial process, are of significant commercial interest.

Thermostable polymer-degrading enzymes such as amylases, pullulanases, xylanases, proteases and cellulases play an increasingly important role in food, chemical, pharmaceutical, paper pulp and waste-treatment industries [Bruins *et al.*, 2001]. The starch-processing industry, for example, is unique within the industrial enzyme sector in that the use of enzymes is essential for the industry. Although the use of enzymes in this industry has resulted in the improvement of both the processes and the products, the operating conditions are limited by the properties of the enzymes used. Hence, recent advances in molecular biology and protein engineering have allowed enzymes with improved operating parameters i.e. increased thermal stability and/or activity, to be introduced to commercial applications resulting in lower operating costs and higher-quality products [Crabb & Mitchinson, 1997].

1.4 Thermostability and thermoactivity

1.4.1 Thermodynamic stability

There is evidence that, for small, monomeric proteins, there is a two-state model for protein unfolding and folding, where the protein exists in only the folded (F) and unfolded (U) forms. Hence in a reversible system:

$$F \underset{k_2}{\overset{k_1}{\rightleftharpoons}} U$$

where k_1 and k_2 are the rate constants for the unfolding and folding reactions, respectively. Thus K_{eq} (the equilibrium constant) = [U]/[F] = k_1/k_2, and ΔG_{FU} (the free energy difference between folded and unfolded forms) is given by :

$$\Delta G_{FU} = -RT \ln K_{eq}$$

where R = the gas constant and T is the absolute temperature.

ΔG_{FU} is a measure of the thermodynamic stability of the protein and can be calculated from the determination of K_{eq} values at varying concentrations of a denaturant such as urea or guanidine hydrochloride, or at varying temperature. For most proteins this value of ΔG_{FU} is small, 20-60kJ/mol, equivalent to a few non-covalent interactions, and thus proteins appear to be only marginally stable at their *in vivo* temperatures of operation. However,

even though ΔG_{FU} is small in magnitude, it is in fact the combination of two large opposing contributions from the enthalpic (ΔH_{FU}) and entropic (ΔS_{FU}) changes between F and U:

$$\Delta G_{FU} = \Delta H_{FU} - T\Delta S_{FU}$$

The predicted contributions of the contact interactions and conformational entropy components of stability in the folding process of a mesophilic protein are shown in Figure 1.1 as a function of temperature. Folding is assumed to be driven principally by hydrophobic interactions and opposed by the conformational entropy [reviewed by Dill *et al.*, 1989; Danson & Hough, 2001: Rees & Robertson, 2001]. The enthalpic (hydrophobic) contribution to ΔG_{FU} strengthens, that is becomes more negative, with increasing temperature, although proteins are denatured at high temperatures despite this. Figure 1.1 predicts two denaturation temperatures, defined by the points at which the $\Delta G_{FU} = 0$ (the temperature at which there are equal quantities of U and F). This arises as the temperature dependence at low temperatures is dominated by the enthalpic changes, the conformational entropy being almost independent of temperature. At high temperatures the conformational entropy dominates the temperature dependence, the slope of the hydrophobic contribution being smaller.

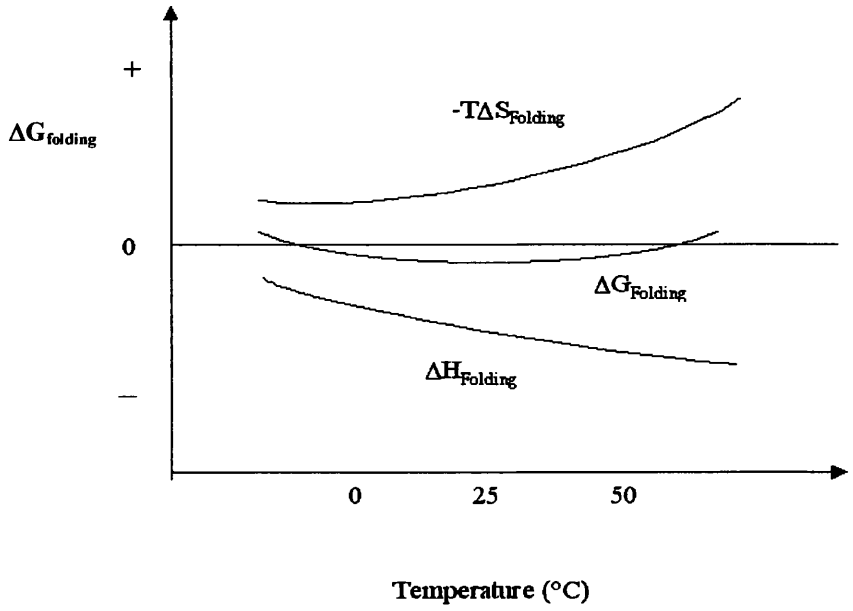


Figure 1.1 **Contributions of enthalpy (ΔH) and conformational entropy ($T\Delta S$) to the free energy of folding of a protein (ΔG).**
A schematic diagram to illustrate how the thermodynamic parameters for protein folding vary with temperature, and to show how the relatively-small value of the free energy of unfolding ($\Delta G_{folding}$) of a mesophilic globular protein comprises relatively-large enthalpic ($\Delta H_{folding}$) and entropic contributions ($-T\Delta S_{folding}$).

Using the data in Figure 1.1, it is predicted that an increased thermostability of a protein must involve a shift in the ΔG_{FU} versus temperature profile compared with its mesophilic counterpart. There are several thermodynamic mechanisms by which the thermostability of proteins could be increased, such as those depicted in Figure 1.2. These possibilities can be considered to arise by a combination of shifting the stability curve, shifting the curve to high temperatures, and broadening the curve [discussed by Danson & Hough, 2001; Rees & Robertson, 2001].

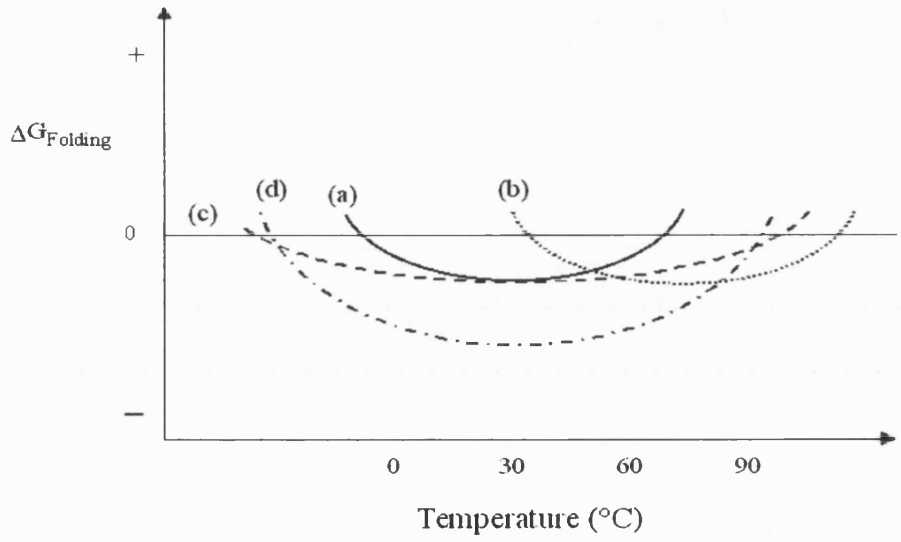


Figure 1.2 **Potential thermodynamic strategies for increasing the thermostability of a protein.**
A typical free energy curve ($\Delta G_{folding}$) for a mesophilic protein is shown (a), as are also the possible strategies for increasing its thermostability by shifting the free energy curve to higher temperatures (b), broadening the curve (c), or increasing the $\Delta G_{folding}$ at all temperatures (d).

1.4.2 Kinetic stability

When measuring the thermostability of an enzyme, a practical consideration is that thermal unfolding is often irreversible, and in such cases the two-state model, which assumes a reversible process as described in Section 1.4.1, must be extended to take account of the irreversible process:

$$F \underset{k_2}{\overset{k_1}{\rightleftharpoons}} U \xrightarrow{k_3} X$$

Where X is the irreversibly-inactivated state and k_3 is the rate constant describing its formation from U. In the case of irreversible unfolding, the measured rate of the overall process from F to X is termed:

$$k_{obs} = \frac{(k_1 \times k_3)}{(k_2 + k_3)}$$

According to this model, the upper limit of k_{obs} is set by k_1 as only U (and not F) will undergo an irreversible transition. As discussed by Sterner *et al.* (2001), this is relevant for thermophilic proteins. At the high physiological temperatures of thermophiles, the overall process will be fast and therefore $k_3 >> k_2$ and hence $k_{obs} \cong k_1$. This means that the rate of irreversible denaturation (k_{obs}) can be measured, and usually its value is much lower than that of the homologous mesophilic proteins. The value of k_1 is determined by ΔG^* (the activation energy from F to U) in contrast to the thermodynamic stability which is determined by ΔG (Figure 1.3).

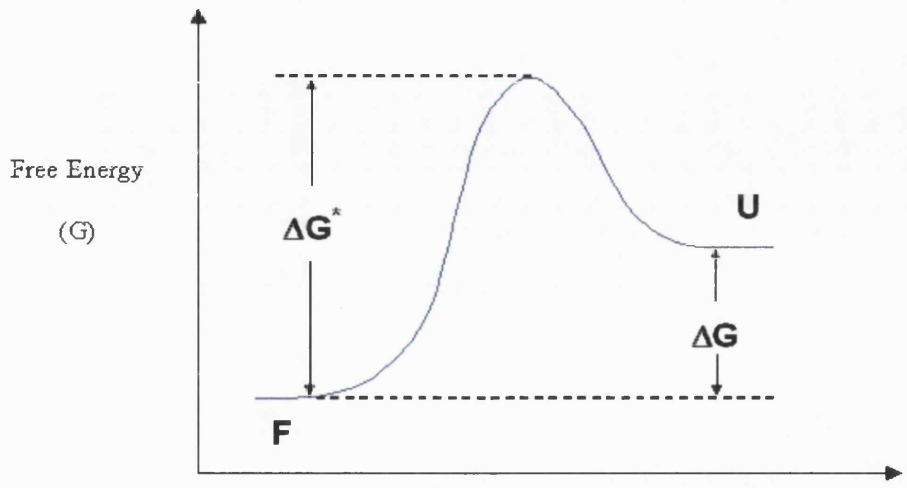


Figure 1.3 **A two-state protein unfolding schematic to illustrate thermodynamic and kinetic stability.**
A schematic diagram to illustrate the change in free energy along the co-ordinate describing the unfolding of a folded protein (F) to its unfolded form (U). ΔG is the free energy change between F and U and is a measure of the thermodynamic stability of the protein. ΔG* is the free energy change between F and the transition state of the process, and is a measure of kinetic stability.

1.4.3 Thermoactivity

Enzymes from thermophilic organisms must not only be stable to high temperatures, but they must also be catalytically active. However, it is important to note that an increased thermal stability does not guarantee thermoactivity.

It is generally observed that, for enzymic reactions, the rate of reaction is increased 2-fold on raising the temperature by 10°C, and from this a more than 60-fold higher rate for the thermoactive enzyme at 100°C than that of the mesophilic at 37°C would be predicted. In the majority of thermophilic enzymes studied, this is not the case as the specific activity of a thermostable enzyme at its optimum temperature is often comparable with that of the homologous enzyme from a mesophile at its optimum [discussed by Danson *et al.*, 1996; Danson & Hough, 2001], suggesting that thermophilic enzymes are poor catalysts. It is clear that molecular flexibility is essential both for substrate binding and catalysis, and enzymes must balance this need for flexibility with the requirements of thermal stability. Thermophilic proteins are thought to be less flexible than mesophilic proteins at mesophilic temperatures [reviewed by Lonhienne *et al.*, 2000; Svingor *et al.*, 2001; Fields, 2001], and it is also been proposed that thermophilic and mesophilic proteins only attain a similar degree of flexibility at their respective optimum temperatures; the comparison of 3-isopropylmalate dehydrogenase (IPMDH) from the thermophilic *Thermus thermophilus* with its mesophilic counterparts illustrates this hypothesis [Závodsky *et al.*, 1998]. Therefore, through the increased rigidity of the structure required for thermal stability,

hyperstable enzymes can only achieve, at their optimum temperature, the same catalytic activity as a mesophilic enzyme does at its optimum temperature [Danson *et al.*, 1996].

1.5 Understanding the structural basis of thermostability

Until recently protein stability studies were restricted to small, soluble, monomeric enzymes such as lysozyme and ribonuclease, and a large proportion of the knowledge of protein-stabilisation mechanisms comes from studying these models [reviewed by Fágáin, 1995]. More recently the importance of studying larger, oligomeric proteins has been realised, as these proteins make up the majority of enzymes, and has led to the investigation of monomer association in the role of protein stability. Having access to thermostable enzymes now allows the study of the extent to which nature uses these stabilisation mechanisms, and also of the additional mechanisms that participate in hyperthermostability

Studying the crystal structures of larger, multimeric homologous enzymes from both mesophilic and thermophilic organisms has revealed that, generally, there is a high degree of structural homology, even if there is low protein sequence homology. This suggests that the thermophilic enzymes rely on the same catalytic mechanisms as do their mesophilic counterparts and that the increased stability of thermostable enzymes, compared with their mesophilic homologues, must be a result of differences in specific amino acid sequences that do not significantly change the overall three-dimensional conformation.

Previous work has shown that heterologous expression of thermostable enzymes in mesophilic hosts generally does not affect the thermal stability and activity of these enzymes, also indicating that, in many thermophilic proteins, thermal stability is an intrinsic property of the protein and does not require intracellular thermoprotectants. Hence, thermophilic and hyperthermophilic enzymes can be expressed and purified easily, and this has allowed the crystal structures of a series of enzymes with increasing thermal stability to be compared, and structural trends to be identified e.g iron superoxide dismutase [Ursby *et al.*, 1999], 3-isopropylmalate dehydrogenase [Nemeth *et al.*, 2000], glutamate dehydrogenase [reviewed by Lebbink *et al.*, 1999], and citrate synthase [Danson & Hough, 1998, 2001]. The structural trends identified to date, are discussed in detail in Section 1.7. In general, it is found that most structural adaptations to thermostability result in a more rigid enzyme.

1.6 Methods of investigating thermostability

As discussed previously, as industry becomes more interested in the use of thermophilic enzymes in commercial processes, there is an increased interest in engineering selected properties into proteins. This section discusses the advantages and disadvantages of the three main methods of investigating protein thermostability: protein sequence comparison, rational design/site-directed mutagenesis, and directed evolution [reviewed by Chen, 2001; Lehmann & Wyss, 2001].

1.6.1 Computer aided sequence comparisons and the consensus concept

Before the wide current availability of protein crystal structures, comparing protein sequences was a valuable method of identifying differences between mesophilic and thermophilic enzyme homologues. Although this method is less informative than structural comparisons, these kinds of data have still highlighted general trends in amino acid propensities. For example, a comparison of the sequences of enzymes from the thermophilic archaeon *Methanococcus jannaschii* and their mesophilic homologues has revealed an increase in residue volume and hydrophobicity, and an increase in the number of charged amino acids [Haney *et al.,* 1999]. Such studies on individual proteins have been supported by the comparative analysis of 20 complete genomes for mesophilic, thermophilic and hyperthermophilic organisms [Thompson & Eisenberg, 1999], showing a trend towards shortened thermophilic proteins relative to their mesophilic homologues; more specifically, thermophilic sequences are more likely to have deletions in exposed loop regions.

One method of determining potential amino acids for mutation to increase thermostability is to compare the amino acid sequences directly, as previously discussed. Liu *et al.* (2000) identified several potential regions of the glucoamylase from *Asperigillus awamori* that may increase thermostability, based on sequence comparisons with other, more thermostable, glucoamylases. They found that all mutations increased glucoamylase thermostability by up to 4°C in operating temperature. However, the many failed attempts

at rational engineering are usually based on primary amino acid sequence homologies as the only criterion for amino acid replacement.

Recently, a semi-rational approach for engineering thermostability was published by Lehmann *et al.* (2000a) [also reviewed by Lehmann *et al.,* 2000b], utilising the computer based comparisons of protein sequences. Interestingly, this method uses the comparison of many homologous mesophilic enzyme sequences. The hypothesis for this method is that at a given position in an amino acid sequence alignment of homologous proteins, the respective consensus amino acid contributes more than average to the stability of the protein than the non-consensus amino acids. Consequently, substitution of nonconsensus by consensus amino acids may be a feasible approach for improving the thermostability of a protein. Hence, this method is called the 'consensus concept'.

Lehmann *et al.* (2000a) compiled a sequence alignment of 13 homologous, mesophilic fungal phytases, which was then used to calculate a consensus phytase amino acid sequence. This consensus protein showed a 15-22°C increase in unfolding temperature compared to each of its parents, but, in spite of the increase in thermostability, catalytic activity at 37°C was not compromised. This may illustrate a link between protein sequence conservation and protein stability i.e. that many mesophiles partially contain the determinants of thermal stability. The main problem with using this method for analysing proteins is the requirement of many homologous protein sequences.

1.6.2 Structural comparison and rational design by site-directed mutagenesis

In the rational design of enzymes, precise changes in amino acid sequence are preconceived based on a detailed knowledge of protein structure, function and mechanism, and are then introduced using site-directed mutagenesis. This technology holds potential for redesigning enzymes for industrial applications, although it is also an important tool for understanding the structural basis of thermostability, substrate binding and catalytic mechanisms of enzymes.

A better strategy for identifying thermostabilising mutations involves the comparison of protein structures of more stable proteins with less stable ones, to identify structural trends with increasing thermostability. The main problem with this method is that the stability of a protein is generally determined by a multitude of local and long-range interactions, discussed in Section 1.7. Hence, in order to achieve increased thermostability, several substitutions, each with a relatively small effect, usually need to be combined in a multiple mutant. Lebbink *et al.* (1999) investigated the role in stability of an 18-residue ion-pair network that is present in the glutamate dehydrogenase (GDH) from the hyperthermophilic Archaeon *Pyrococcus furiosus*, by introducing four new charged amino acids into the subunit interface of the homologous enzyme from the thermophilic bacterium *Thermotoga maritima*. The *T.maritima* GDH has a less extensive ionic network compared to the *P.furiosus* GDH. Amino acid substitutions were introduced as single mutations as well as in combination. They showed that many of the single mutations resulted in a decrease in thermal stability, but a combination of destabilising single mutations, in most cases,

restored or increased thermal stability i.e. an increase in temperature optima of 3°C, and a 0.5°C higher melting temperature. This indicates the need for multiple small changes to achieve pronounced thermostabilisation.

In the investigation of the structural basis of thermostability, rational design/site-directed mutagenesis has identified many amino acids involved in achieving a thermally stable enzyme by either removing the amino acid and showing there is a decrease in thermal stability, or introducing a residue and showing an increase in thermostability. A reduction in thermal stability has been shown in the thermostable D-glyceraldehyde-3-phosphate dehydrogenase [Pappenberger *et al.*, 1997], indoleglycerol phosphate synthase [Merz *et al.*, 1999] and citrate synthase [Arnott *et al.*, 2000], by removing residues essential for thermostability. More importantly, it is also possible to increase the thermostability of a mesophilic or moderately thermophilic enzyme by the addition of specific residues as demonstrated for a moderately stable protease [Van den Burg *et al.*, 1998], the thermostable citrate synthase from *T.acidophilum* [Erduran & Kocabıyık, 1998] and the mesophilic Lac Repressor from *E.coli* [Gerk *et al.*, 2000].

A major drawback of this method of rational protein redesign is that, since it relies on the sequence or structural comparisons of homologous enzymes, there is the problem of identifying the relevant thermostabilising mutations in less homologous or highly divergent sets of proteins. Also, after PCR, it requires confirmation of the mutation by sequencing and then by purification of the mutant enzymes following each round of mutagenesis in

order that kinetic and functional properties can be determined. This is time consuming and expensive, and may be impractical for multiple cycles of mutagenesis.

1.6.3 Directed evolution

A major advantage of directed evolution over rational design/site-directed mutagenesis is that the former technique does not require information about how enzyme structure relates to function. Directed evolution employs a random process in which multiple rounds of error-prone PCR are used to create a library of mutated genes from a parent gene. After each round, the best mutants are selected and used as parent sequences in the following round of random mutagenesis. This technique has been recently reviewed by Arnold *et al.* (2001); Chen (2001) and Lehmann & Wyss (2001). Recently, directed evolution has been significantly improved using *in vitro* recombination or DNA shuffling [reviewed by Moore *et al.*, 1997]. Genetic selection or high-throughput screening subsequently identifies the mutants that possess the required improved properties. However, a major limitation of this method is the prerequisite for a sensitive and efficient method for screening a large number of potential mutants.

The increases in thermostability obtained by directed evolution have, as yet, been no more impressive than the best examples of rational design, although they may have required less time and effort [referenced by Lehmann & Wyss, 2001]. However, more importantly, directed evolution may generate new variants of an enzyme that are not encountered in proteins from naturally existing organisms, but are optimal for industrial applications

[Arnold & Volkov, 1999]. For this reason, directed evolution is a valuable tool widely adopted by industry in the search for novel enzymes with biotechnological application e.g. enzymes active in solvents, enzymes with altered substrate specificity or enzymes with increased thermostability [reviewed by Petrounia & Arnold, 2000]. Miyazaki *et al.* (2000) found that after a heat-sensitive psychrophilic protease subtilisin S41 was subjected to three rounds of mutagenesis/recombination and screening, the resulting variant had a half-life at 60°C that was approximately 500 times that of the wildtype. Interestingly, the evolved protease subtilisin retained its activity at low temperature in spite of its increased thermostability. Wintrode *et al.* (2001) subjected this evolved subtilisin to five additional generations of laboratory evolution and reported a 1200 times increase in half-life at 60°C compared to the wild-type enzyme.

Another valuable use of the directed evolution method is to study the natural adaptation of enzymes for stability and function [reviewed by Arnold *et al.*, 2001]. As previously described, many extremophilic enzymes are similar to their counterparts from mesophilic environments. Sequence comparisons suggest that these enzymes are derived from a common ancestral enzyme and have accumulated mutations that allow them to adapt over millions of years. An understanding of these mechanisms of adaptation might allow the production of enzymes with enhanced properties for industrial processes.

The sequence comparison/consensus concept, rational design and directed evolution have their distinct advantages and in the case of engineering proteins for thermostability, there is the choice between three different, apparently equally successful, strategies. The use of the

'consensus concept' as a tool for the construction of a protein with an increased thermal stability, compared with rational design and directed evolution, is discussed by Lehmann & Wyss (2001). Not only have these techniques allowed the production of enzymes with increased thermostability, they have been excellent tools for the study of the structural basis of thermostability. Section 1.7 discusses the general features of thermostability identified using these techniques.

1.7 General intrinsic features of thermostability

There have been many recent reviews summarising the current understanding of the structural basis of protein thermostability. Three main reviews were referred to in the following discussion of the general intrinsic features of thermostability: Vieille & Zeikus (2001), Danson & Hough (2001) and Sterner & Liebl (2001).

1.7.1 Conformational flexibility and loop regions

As discussed in Section 1.4, a reduction in the conformational entropy of the unfolded state of the protein will reduce the value of ΔS_{FU} and consequently will promote folding. This reduction in conformational entropy (reduction in flexibility) of the unfolded state of thermally stable enzymes may be achieved in numerous ways. Firstly, in the unfolded state, glycine is the residue with the highest conformational entropy and proline, which adopts only a few configurations, has the lowest entropy. These features were utilised by Van den Burg *et al.* (1998) to engineer an enzyme to resist boiling. Rigidifying mutations

such as Gly to Ala and Ala to Pro were introduced into the moderately stable, thermolysin-like protease from *Bacillus stearothermophilus*. Whereas the wild-type enzyme had an undetectable short half-life at 100°C, the mutant had a half-life of 170 min at this temperature. The role of proline substitutions in protein thermostabilisation has been reviewed by Watanabe & Suzuki (1998).

Secondly, a comparison of 20 genomic sequences showed that thermophilic proteins are shorter than their mesophilic homologues and, where atomic structures were available, the shortening was found to be mainly achieved within the loop regions of the protein and at the N- and C-termini [Thompson & Eisenberg, 1999]. Through their inherent flexibility, loops are considered to be potential initiation points for thermal denaturation, and therefore a reduction in their size (and hence flexibility) may lower the conformational entropy of the folded protein. A reduction in the conformational entropy of the folded state results in an increased rigidity of the folded enzyme, which is a common feature of thermostability, although this may reduce the comparative catalytic efficiency of the thermophilic enzyme at mesophilic operation temperatures (see Section 1.4.3). The reduction of loop regions as a feature of thermostability is discussed with reference to thermostable citrate synthases (CSs) in Section 1.12.2.

Shortening of the polypeptide chain, by reduction of loop regions, contributes to the thermal stability of the tryptophan synthase α-subunit from the hyperthermophile *Pyrococcus furiosus* [Yamagata *et al.*, 2001]. A decrease in cavity volume and an increase

in ion pairs also contributed in the thermal stability of this protein, and these features are discussed later.

1.7.2 Increased hydrophobicity

An increase in the hydrophobic interactions can increase the thermostability of a protein. Increased burial of hydrophobic residues into the interior of the protein will be entropically favourable in terms of the lower degree of ordering of water molecules that occurs when these side chains are removed from the solvent. Hence, hydrophobic interactions are often thought to provide the energy required for proteins to fold in aqueous solutions.

At mesophilic temperatures, transfer of a non-polar molecule from an aqueous environment to a non-polar environment results in the release of water molecules tightly ordered around non-polar residues. An unfavourable enthalpic contribution due to the disturbance of the water molecules is compensated by the favourable enthalpic contributions from van der Waals interactions. This results in a net enthalpy change of zero. Hence, at these temperatures the formation of hydrophobic interactions is favourably driven by the change in entropy when the ordered water molecules are removed. At higher temperatures, the entropic contribution is less favourable as the water molecules are initially less ordered around the non-polar molecules due to the decrease in the dielectric constant of water at these temperatures. As there are no unfavourable enthalpic contributions due to the disturbance of ordered water molecules, the formation of hydrophobic interactions at higher temperatures is enthalpy driven. This phenomenon, known as the 'entropy-enthalpy

compensation', means that the free-energy change of transferring a non-polar molecule from an aqueous to a non-polar environment alters less with temperature than do either the entropy or the enthalpy, although the net result is that the hydrophobic effect strengthens, at least in the 0-80°C range. Moreover, because the free-energy change is small, this increase in hydrophobic interactions at higher temperatures may make a significant contribution to the thermostabilisation of thermophilic proteins.

Above 80°C, the hydrophobic interaction strengthens by smaller and smaller increments with a given rise in temperature, and reaches a maximum around 100-140°C, when the entropy of transfer becomes zero; thereafter it decreases in strength. Enzyme stability in this region may thus require additional interaction, and this may explain the prevalence of ionic bonding in hyperthermophilic proteins [Danson & Hough, 2001].

Hydrophobic interactions have been predicted to play a role in protein stability from a comparison of the amino acid sequences of thermophilic enzymes and their mesophilic counterparts [Haney *et al.*, 1999] and can contribute to the improved packing of the protein [Wallace *et al.*, 2000]. Studies on the Lac repressor of the mesophilic *Esherichia coli* [Gerk *et al.*, 2000] and the citrate synthase from the thermophilic *Thermoplasma acidophilum* [Erduran & Kocabıyık, 1998] have shown that an increase in hydrophobic interactions at the subunit interface can increase the temperature optimum (optimal substrate binding) by up to 40 °C, in the case of the Lac repressor. Although the increase in hydrophobic residues seems to contribute to the thermostability of thermophilic proteins, it has been noted that the role of hydrophobic interactions is less in hyperthermophilic proteins, and

ionic interactions seem to contribute more to hyperthermostability [Szilágyi & Závodsky, 2000], as mentioned before. The role of ionic interactions in thermostability is discussed in the next section.

1.7.3 Ionic interactions and networks

Coulomb's law states that the energy of the electrostatic interaction between two atoms A and B, in a vacuum, is simply the product of their charges divided by the distance between them (r_{AB}):

$$\Delta E = \frac{Z_A Z_B \in^2}{r_{AB}}$$

where $\in$ is the charge of the electron and Z the number of such charges on each atom. If the two charges are of opposite sign, the energy (ΔE) decreases as they approach each other, and the interaction is favourable; if the charges are of the same sign, there is repulsion between them.

Coulomb's law is valid only for two point charges in a vacuum. For other environments, the electrostatic interaction is modulated by other interactions. In homogenous environments, the electrostatic interaction is decreased by the dielectric constant, so the electrostatic interaction in media other than a vacuum is always less than that stated by Coulomb's law. Liquids have dielectric constants in the range of 20-110, and water has a dielectric constant of ~80.

As previously mentioned, Szilágyi and Závodsky (2000) compared structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits. The only general rule they observed was an increase in the number of ion pairs with increasing growth temperature. Genomic comparisons show that proteins from thermophilic, and more specifically hyperthermophilic, organisms tend to have a higher percentage of the charged amino acids, Glu, Lys and Arg, than those of mesophiles [Cambillau & Claverie, 2000]. This is consistent with the observed increase in the number of ionic interactions in some thermostable proteins [Vogt *et al.,* 1997; Cavagnero *et al.,* 1998; Kumar *et al.*, 2000a; 2000b; Szilágyi & Závodsky., 2000].

Salt bridges have been proposed to play a crucial role in promoting hyperthermostability in proteins, yet experimental and theoretical studies have indicated that at room temperature salt bridges generally destabilise, or at best only slightly stabilise the folded state of proteins [reviewed by Elcock, 1998]. This has also been recently reviewed by Danson & Hough (2001) and Karshikoff & Ladenstein (2001).

The equation below describes the energetics of the formation of an ionic bond in a protein:

$$\Delta G_{total} = \Delta G_{desolvation} + \Delta G_{electrostatic} + \Delta G_{protein}$$

where ΔG_{total} is the overall free-energy of the ionic bond (the formation of an ionic bond requiring a 'favourable', or negative value, ΔG_{total}) and is made up of three primary components:

1. The $\Delta G_{desolvation}$ describes the energy required to remove the water molecules from the charged residues, this is always unfavourable (positive) at room temperature and, hence, is often called the 'desolvation penalty'.

2. The $\Delta G_{electrostatic}$ describes the favourable van der Waals and Coulombic interactions associated with the formation of an ionic bond.

3. The $\Delta G_{protein}$ describes the favourable interactions of ionic bonding residues with other charged amino acids in the three-dimensional structure of the protein.

At lower temperatures the $\Delta G_{electrostatic}$ and the $\Delta G_{protein}$ contributions are favourable for the formation of an ionic bond but, as the desolvation penalty is very unfavourable (positive), the resulting ΔG_{total} is unfavourable. At higher temperatures the dielectric constant of water reduces, whereas the dielectric constant of the protein is insignificantly changed, and the desolvation energy becomes less unfavourable i.e. at high temperatures water molecules have a greater mobility and, therefore, they hydrate a charge less well than at lower temperatures. Hence, an increase in temperature results in the $\Delta G_{desolvation}$ becoming less unfavourable, and the $\Delta G_{electrostatic}$ and the $\Delta G_{protein}$ becoming more favourable, which results in an overall favourable ΔG_{total}.

An interesting observation is that there is an increase in the number of ionic networks in hyperthermophilic proteins, particularly at subunit interface regions; this is discussed later in Section 1.7.5. Several enzymes from the hyperthermophilic Archaeon *Pyrococcus furiosus* (growth optima ~100°C) have been well studied, most specifically citrate synthase and glutamate dehydrogenase. Analysis of the crystal structures of both citrate synthase [Russell *et al.*, 1998] and glutamate dehydrogenase [Yip *et al.*, 1995] from *Pyrococcus furiosus* has revealed extensive ionic networks at subunit interface regions that are absent in their mesophilic counterparts. This may suggest that , although there is a general increase in ion pairs in thermophilic proteins, the presence of ionic networks may be a feature of hyperthermophilic proteins. The ionic networks present in the *P.furiosus* citrate synthase are discussed in detail later.

1.7.4 Packing density

The overall compactness of a protein can be reflected by a low surface area to volume ratio, increased packing density, fewer solvent exposed residues and internal cavities, and a shortening of loop regions [discussed by Bell, 1999], which can contribute to an overall increase in rigidity.

Also, van der Waals forces between hydrophobic groups buried within the interior of the protein can result in an enthalpic contribution to the strength of the hydrophobic interaction. Hence, optimised packing within a protein might lead to better van der Waals interactions and hence increased thermostability. Chen *et al.* (2000) constructed seven hyper-stable

multiple mutants of staphylococcal nuclease and found that their thermal denaturation midpoint temperatures were 12.6 to 22.9°C higher than that of the wild-type protein. Detailed structural analysis of the crystal structures of these mutants showed that packing had improved, with the volume of the mutant proteins' hydrophobic core decreasing as protein stability increased. Further, the number of van der Waals interactions in the entire protein showed an experimentally significant increase correlated with increasing stability.

1.7.5 Oligomeric enzymes and subunit interactions

Approximately 80% of globular proteins, and hence the vast majority of enzymes, are oligomeric, and it is becoming increasingly common to find that their active sites are situated between component subunits, with residues from different monomers contributing to the active site. Therefore, it is crucial for thermophilic enzymes to maintain the integrity of their oligomeric structure at high temperatures.

Recent studies have suggested that the integrity of the oligomeric protein structure in thermophilic enzymes is maintained by hydrophobic interactions; for example, citrate synthase from the thermophilic *Thermoplasma acidophilum* [Russell *et al.*, 1994]. Erduran & Kocabıyık (1998) showed that increasing the number of hydrophobic interactions at the subunit interface of citrate synthase from *T.acidophilum* enhanced the half-life of the enzyme up to 23-fold at 85°C compared to the wild-type enzyme. Increasing the hydrophobic interactions at the subunit interface of the Lac repressor from a mesophilic

organsism also increased the temperature at which substrate binding is retained by 40°C compared to the wild-type [Gerk *et al.*, 2000].

Comparisons of crystal structures of mesophilic, thermophilic and hyperthermophilic proteins have also suggested that the inter-subunit contact regions in hyperthermophiles are stabilised by ionic interactions [Rahman *et al.*, 1998; Russell *et al.*, 1997]. Not only is the number of charged residues often increased at the subunit interface of hyperthermophilic organisms, but there have been many cases reported where the charged residues are organised into an ionic network e.g. glutamate dehydrogenase from the hyperthermophilic Archaeon *Pyrococcus furiosus* [Yip *et al.*, 1995], citrate synthase from *Pyrococcus furiosus* [Russell *et al.*, 1997], and D-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Thermotoga maritima* [Pappenberger *et al.*, 1997]. Disruption of these ionic networks has confirmed their involvement in the hyperthermostability of proteins [Pappenberger *et al.*, 1997; Arnott *et al.*, 2000]. There have also been attempts to engineer ionic networks into the subunit interface regions of less thermostable proteins to increase their thermostability [Lebbink *et al.*, 1999], and this will be discussed in detail later.

1.7.6 Thermolabile amino acids

A number of amino acids are chemically degraded in free solution at the high temperatures (>85°C) experienced by hyperthermophilic organisms. Of the 20 amino acid residues, four (Asn, Gln, Met, and Cys) may be classified as thermolabile due to their tendency to

undergo deamidation or oxidation at high temperatures. Many studies involving the comparison of thermophilic and mesophilic protein sequences have shown that there are significant differences in the propensities of certain amino acids in thermostable proteins. Sriprapundh *et al.* (2000) showed that a thermostable xylose isomerase contained a few additional Pro residues and fewer Asn and Gly residues. Comparison of protein sequences from mesophilic and extremely thermophilic *Methanococcus* species showed an increase in charged amino acids (Glu, Arg, and Lys) and a decrease in polar residues (Ser, Thr, Asn, and Gln) [Haney *et al.*, 1999]. Interestingly, there is no observed increase in Asp residues with increasing thermal stability, even though it is a charged residue. This may be because Asp residues are known to form succinimide intermediates at high temperature, which may be hydrolysed to yield a break in the peptide chain [Daniel *et al.*, 1996].

1.8 Extrinsic features of thermostability

All of the previously described thermostabilising features are of an intrinsic nature. However, it has been observed that some extrinsic factors play a role in protein stability *in vivo*; for example substrate and salt concentrations, organic solute concentrations, and pressure (the pressure-induced stabilisation of glutamate dehydrogenase from *P.furiosus* is discussed by Sun *et al.* (1999)). These features are summarised by Vieille & Zeikus (2001). Martins *et al.* (1997) examined the accumulation of organic solutes under optimum growth conditions in twelve species of thermophilic and hyperthermophilic Archaea, and found that there was a correlation between a raise in growth temperature and the accumulation of certain solutes e.g. an increase in the concentration of di-myo-inositol-1,-

1'(3,3')-phosphate with the growth of *P.furiosus*, suggesting that these compounds serve as general stress solutes.

Many features have been identified with increasing thermostability, but it also clear that the mechanism of thermal stability varies between proteins and seems to be the accumulation of several features. Here, glutamate dehydrogenase is discussed as a case study for the investigation of the structural basis of thermostability.

1.9 Glutamate dehydrogenase as a model enzyme for the study of thermostability

Among the enzymes used a models for investigating adaptation mechanisms, one of the best-studied is glutamate dehydrogenase (GDH) [reviewed by Lebbink *et al.*, 1999]. GDH catalyses the oxidative deamination of glutamate to 2-oxoglutarate with concomitant reduction of $NAD(P)^+$. A large number of primary sequences and several three-dimensional structures from GDHs obtained from organisms with a wide range of optimal growth temperatures are available. Hence, the crystal structures of GDHs spanning the biological temperature range can be compared to identify structural trends with increasing thermostability.

In general, GDH is a multimeric enzyme consisting of six identical subunits. The hexamer is composed of two trimers that are stacked upside down on top of each other (Figure 1.4). Each subunit contains an N-terminal, substrate-binding domain that forms almost all of the

inter-subunit interactions, and a C-terminal domain binding the cofactor $NAD(P)^+$ [reviewed by Lebbink *et al.*, 1999].

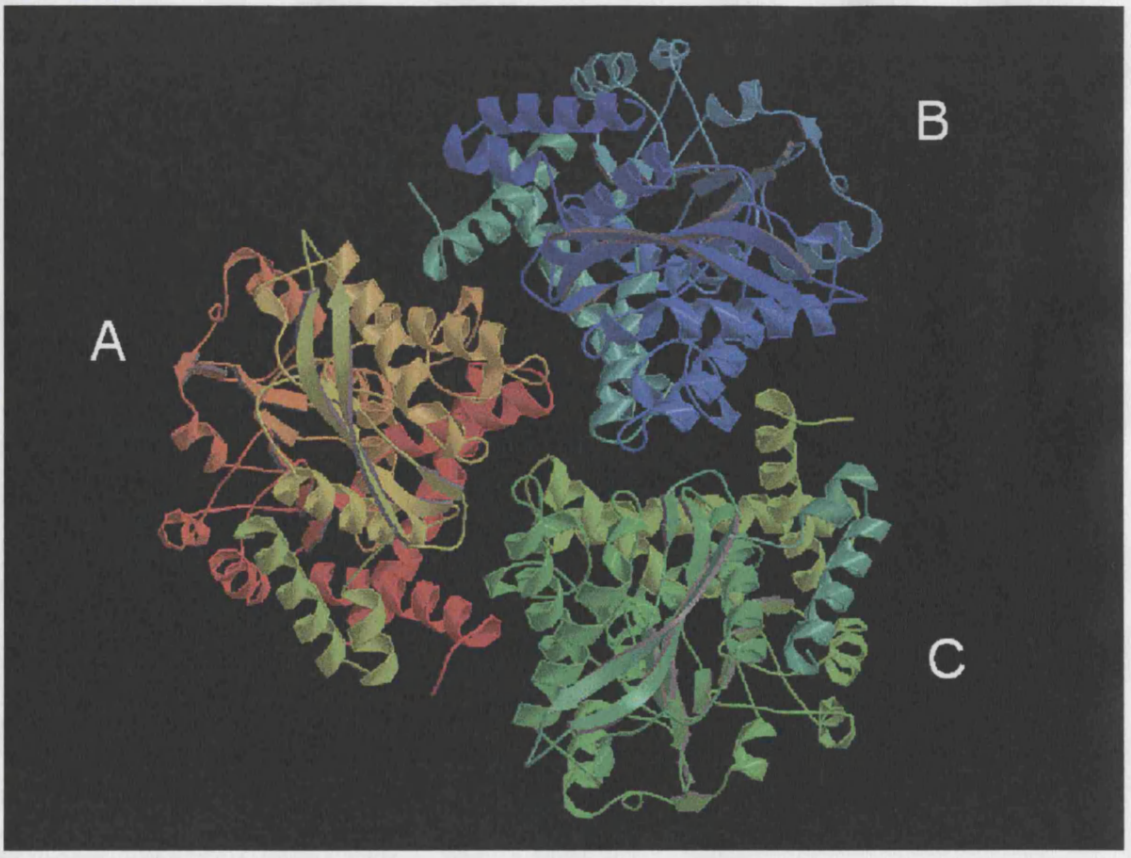


Figure 1.4 **The three-dimensional structure of glutamate dehydrogenase (GDH) from the hyperthermophilic Archaeon *Pyrococcus furiosus*.**
The structure is formed from association of two trimers that are stacked on top of each other. This figure shows the structure of the trimers, with each monomer of the trimer associated with its opposing monomer, labelled A, B, and C.

The determination of the three-dimensional structure of the GDH from the hyperthermophilic archaeon *Pyrococcus furiosus* has allowed the comparison of this structure with those of mesophilic GDHs. The crystal structure of the extremely

thermostable GDH from *Thermococcus litoralis* has been compared with that from *P.furiosus* [Britton *et al.*, 1999]. The two enzymes are 87% identical in sequence, yet differ 16-fold in their half-lives at 104°C. The major structural differences identified from this comparison were that the less stable *T.litoralis* GDH has a decreased number of ion pair interactions, modified patterns of hydrogen bonding and substitutions that decrease packing efficiency. As previously discussed, these are common structural trends identified with a decrease in thermostability.

The most interesting structural feature identified with increasing thermostability of GDHs is the increase in numbers of ion pairs and the formation of ion-pair networks [reviewed by Yip *et al.*, 1998]. As discussed in Section 1.7.3, comparison of the *P.furiosus* GDH with mesophilic GDHs, namely from *Clostridium symbiosum* and *Esherichia coli*, has highlighted the formation of extended networks of ion-pairs as a possible explanation for the thermal stability of the hyperthermostable enzyme. Much work has been carried out in attempts to analyse the extent of formation of ion-pair networks in the different GDHs, and to correlate this with observed thermal stability.

Yip *et al.* (1998) compared both the amino acid sequences and three-dimensional structures of GDHs spanning the biological temperature range. They found that the *P.furiosus* GDH contained three unique ion-pair networks: a large 18-residue ionic network at the N-terminal domains, a 6-residue ionic network and a smaller 3-residue network. The large 18-residue ionic network is formed by the interaction of residues from four monomers across the two-fold axis relating dimers (Figure 1.5). These results indicate that the ion-

pair networks become more fragmented as the temperature stability of the enzyme decreases. Kumar *et al.* (2000a) also compared crystal structures of GDHs to investigate the location, statistics and electrostatic strengths of ion-pairs and of their networks within corresponding monomers of the thermophilic and mesophilic enzymes. They found that the ionic bonds in *P.furiosus* GDH are highly stabilising and concluded that ion-pairs and ionic networks have an important role in resisting unfolding of the protein structure at high temperatures.

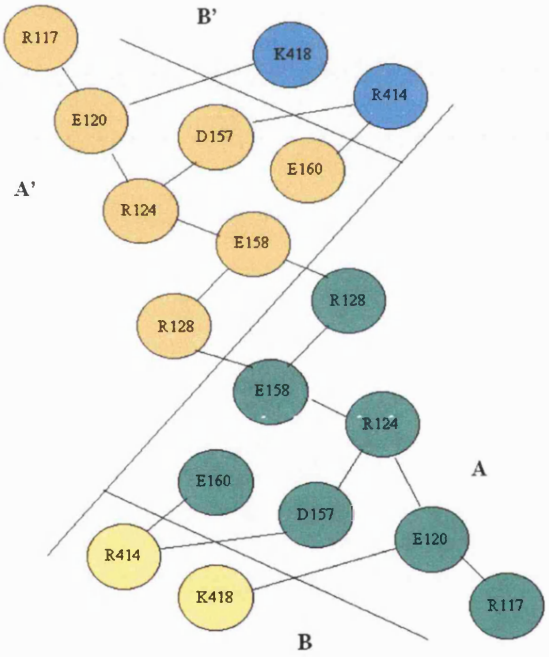


Figure 1.5 **Schematic of the 18-membered ionic network present at the dimer interface of glutamate dehydrogenase from *Pyrococcus furiosus*.** Diagram showing the ion-pair network involving two-fold symmetric interactions of 18 charged residues, based on a figure by Karshikoff & Ladenstein (2001).

Several groups have used the information obtained from sequence and structural comparisons of the GDHs to study the roles of ionic networks in GDH by rational design/site-directed mutagenesis. Rahman *et al.* (1998) attempted to add and remove ion-pairs at the intersubunit of GDH from another hyperthermophilic *Pyrococcus* species, *Pyrococcus kodakaraenis.* They showed that removal of an ion-pair resulted in a lower optimum temperature and a decrease in thermal stability compared to the wild-type GDH, yet the addition of an ion-pair increased the optimum temperature (by approximately 5°C) and the thermostability (a 1h increase in half-life at 100°C).

Lebbink *et al.* (1999) attempted to increase the thermostability of the GDH from the hyperthermophilic bacterium *Thermotoga maritima* by mirroring the 18-residue ionic network found in the *P.furiosus* GDH, as previously discussed in Section 1.6.2. They found that a combination of three destabilising mutations resulted in an enzyme with a 30-minute longer half-life at 85°C, a 3°C higher temperature optimum for catalysis, and a 0.5°C higher apparent melting temperature.

These studies indicate that subunit interactions, specifically extensive ionic networks, are important in the hyperthermostability of GDHs, and may be a general feature of thermostable enzymes.

1.10 Citrate synthase as a model enzyme for the study of thermostability

Citrate synthase is found throughout Bacteria, Eukarya and Archaea, and is chosen here as a model system in which to study the structural basis of extreme stability of thermostable proteins. This enzyme is an ideal model protein as it is present in all phylogenetic domains, there is an extensive data-base of enzymic and structural information from mesophiles, and more recently, thermophiles, and it is an oligomeric protein for which the catalytic mechanism is known in great detail.

1.10.1 Metabolic role

Citrate synthase is a member of a small class of enzymes that can directly form a carbon-carbon bond in the absence of metal ion cofactors. After acetyl coenzyme A (AcCoA) is formed from pyruvate, citrate synthase catalyses the condensation of oxaloacetate (OAA) and AcCoA to form citrate and coenzyme A via a citryl-CoA intermediate (Figure 1.6). The enzyme catalyses the condensation of acetyl-CoA and OAA as the first step in the citric acid cycle, therefore controlling the entry of carbon into the citric acid cycle. Hence, this enzyme occupies a key position in the central metabolism throughout the majority of Bacteria, Eukarya and Archaea.

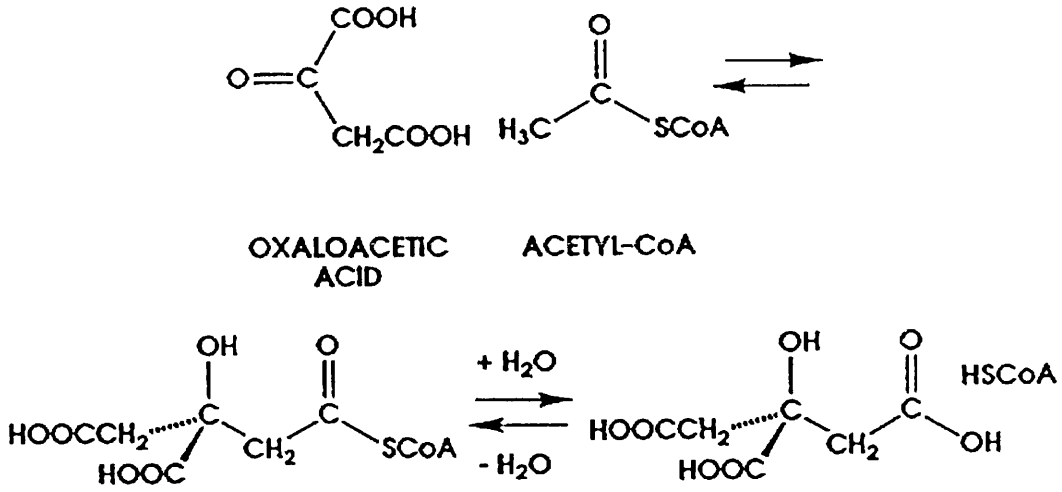


Figure 1.6 The reaction catalysed by citrate synthase [Remington, 1992]

1.10.1 Structure

Analysis of the sizes of citrate synthases from various sources revealed that they can be divided into two groups. The enzyme found in Eukarya, gram-positve Bacteria and Archaea is homo-dimeric in structure, whereas most Gram-negative bacteria contain a citrate synthase that is hexameric.

Study of the three-dimensional structures of mesophilic citrate synthase from both pig (pigCS) and chicken heart has shown that, in eukaryotes, the enzyme is a dimer in which each subunit can be described as consisting of two domains, large and small [Remington, 1982; reviewed by Remington, 1992]. Figure 1.7 shows the three-dimensional crystal

structure of pigCS. The Remington study also revealed that there is a large conformational change induced by substrate binding, with the shift from the open (ligand-free) to the closed (ligand-bound) form being described by an 18° rotation of the small domain relative to the large domain.

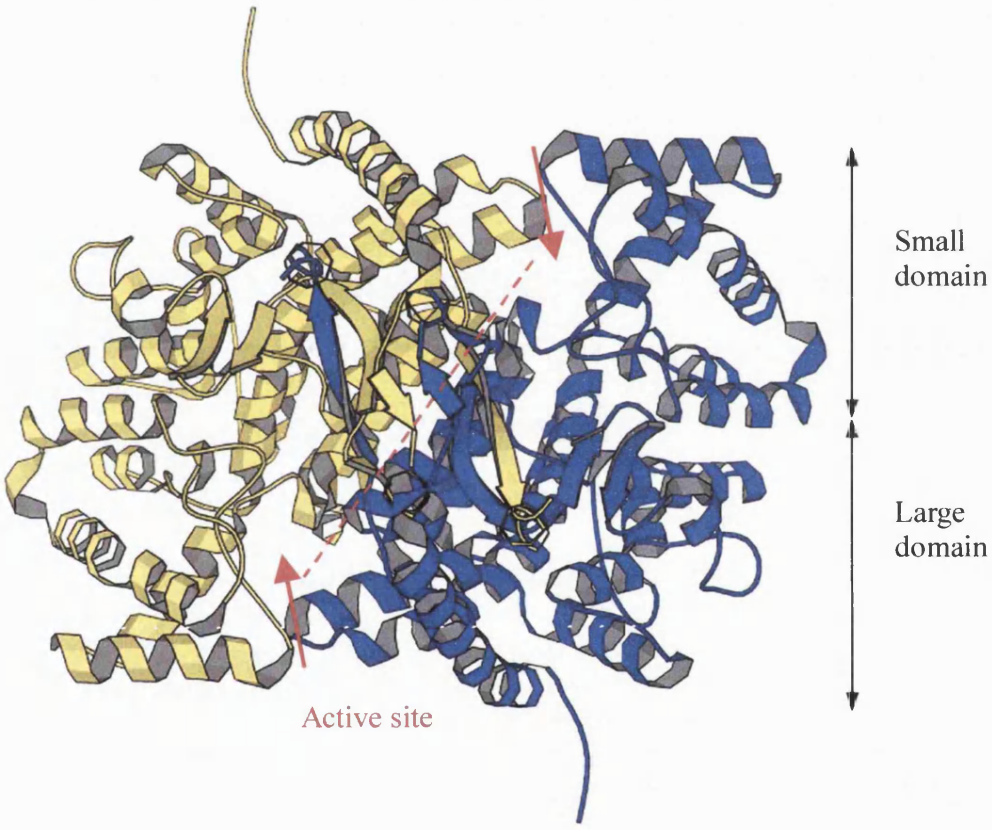


Figure 1.7 **A ribbon diagram of the three-dimensional structure of citrate synthase from pig.** In this figure, one monomer is shown in blue and the opposing monomer is shown in yellow. The subunits associate via the large domains, as shown by the dashed line. The active sites lie between the large and small domains, requiring residues from both monomers, as indicated by the red arrows.

1.10.3 Reaction Mechanism

The citrate synthase reaction is thought to proceed via an acid/base mechanism [Remington, 1992]. It does this by an initial enolisation (deprotonating the methyl group of the acetyl-CoA) followed by a condensation reaction in which the carbanion attacks the carbonyl group of OAA, producing citryl-CoA (Figure 1.8). The final stage involves a thioester hydrolysis resulting in product formation.

This catalytic mechanism was studied initially by determining the three-dimensional crystal structures of substrate or transition-state analogue bound enzymes [Remington, 1982; Karpusas *et al.*, 1990]. These studies identified the potential amino acids involved in substrate binding and catalysis, and their proposed roles were then confirmed using site-directed mutagenesis. The citrate synthase reaction has also been modelled using combined quantum mechanical/molecular mechanical (QM/MM) potential functions [Mulholland & Richards., 1998a; 1998b], and these studies confirmed the findings of the site-directed mutagenesis.

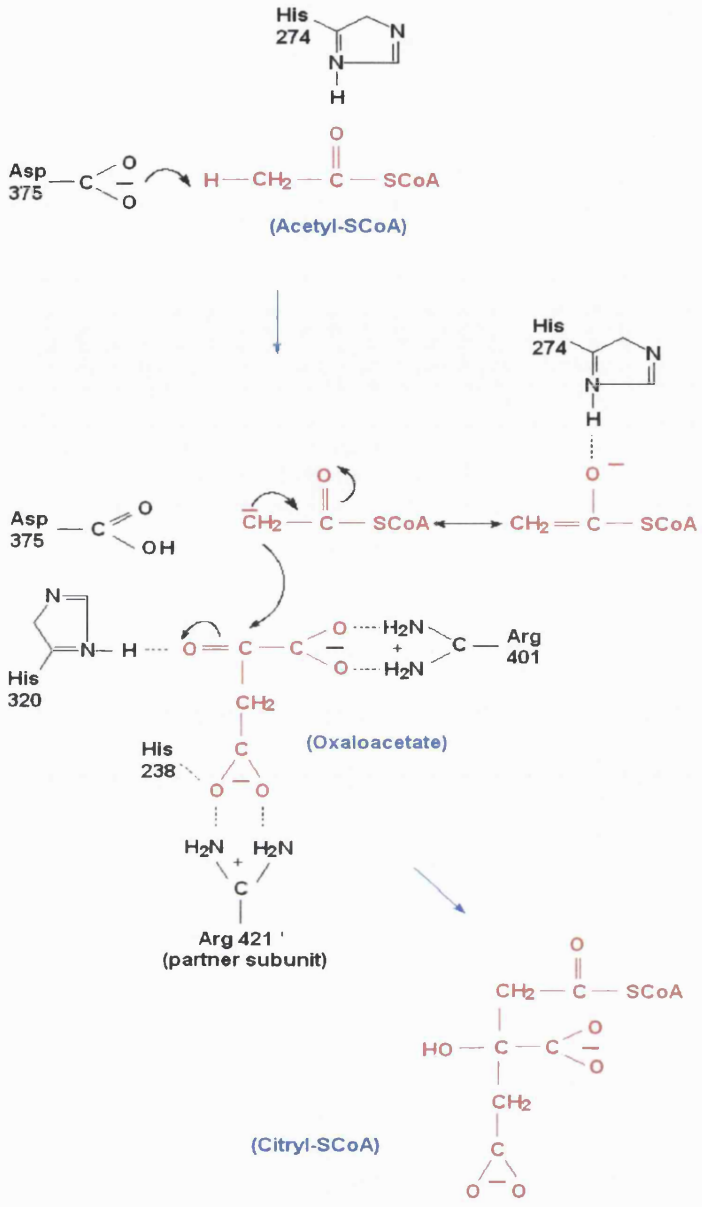


Figure 1.8 **Catalytic mechanism of the condensation reaction catalysed by citrate synthase.**
This schematic diagram describes the condensation of acetyl-CoA and oxaloacetate to citryl-SCoA catalysed by citrate synthase. The residues currently thought to be involved in the mechanism, both catalytic and substrate binding, are shown. These residues are numbered according to the pig citrate synthase. The substrates and products are shown in red.

Many studies of the catalytic mechanisms of the citrate synthase enzyme have involved the site-directed mutagenesis of the catalytic residues His274, His320 and Asp375, as mentioned previously. Initial site-directed mutagenesis studies focused on the substitution of His 274 and Asp375 in the pig heart citrate synthase [Evans *et al.*, 1989; Alter *et al.*, 1990; Zhi *et al.*, 1991] which demonstrated the essential role of these residues in the catalytic mechanism. Recent work has also included the site-directed mutagenesis of His320 in pig citrate synthase [Kurz *et al.*, 1995; 1998; Evans *et al.*, 1996] and the equivalent residues in *E.coli* [Man *et al.*, 1994], which again demonstrates the importance of these conserved residues in catalysis.

1.11 Thermostable citrate synthases

Bacterial expression systems have been developed for the over-expression of citrate synthase enzymes from a number of thermophilic and hyperthermophilic Archaea i.e. *Thermoplasma acidophilum, Ta*CS (growth optimum ~55°C) [Sutherland *et al.*, 1990], *Sulfolobus solfataricus, Ss*CS (~87°C) [Connaris *et al.*, 1998] and *Pyrococcus furiosus, Pf*CS (~100°C) [Muir *et al.*, 1995]. Kinetic and thermal properties of these recombinant enzymes are comparable to the properties of the native citrate synthase enzymes [Danson *et al.*, 1985; Smith *et al.*, 1987]. Sequence comparisons show a high degree of sequence identity, especially between the thermophilic citrate synthases (Figure 1.9).

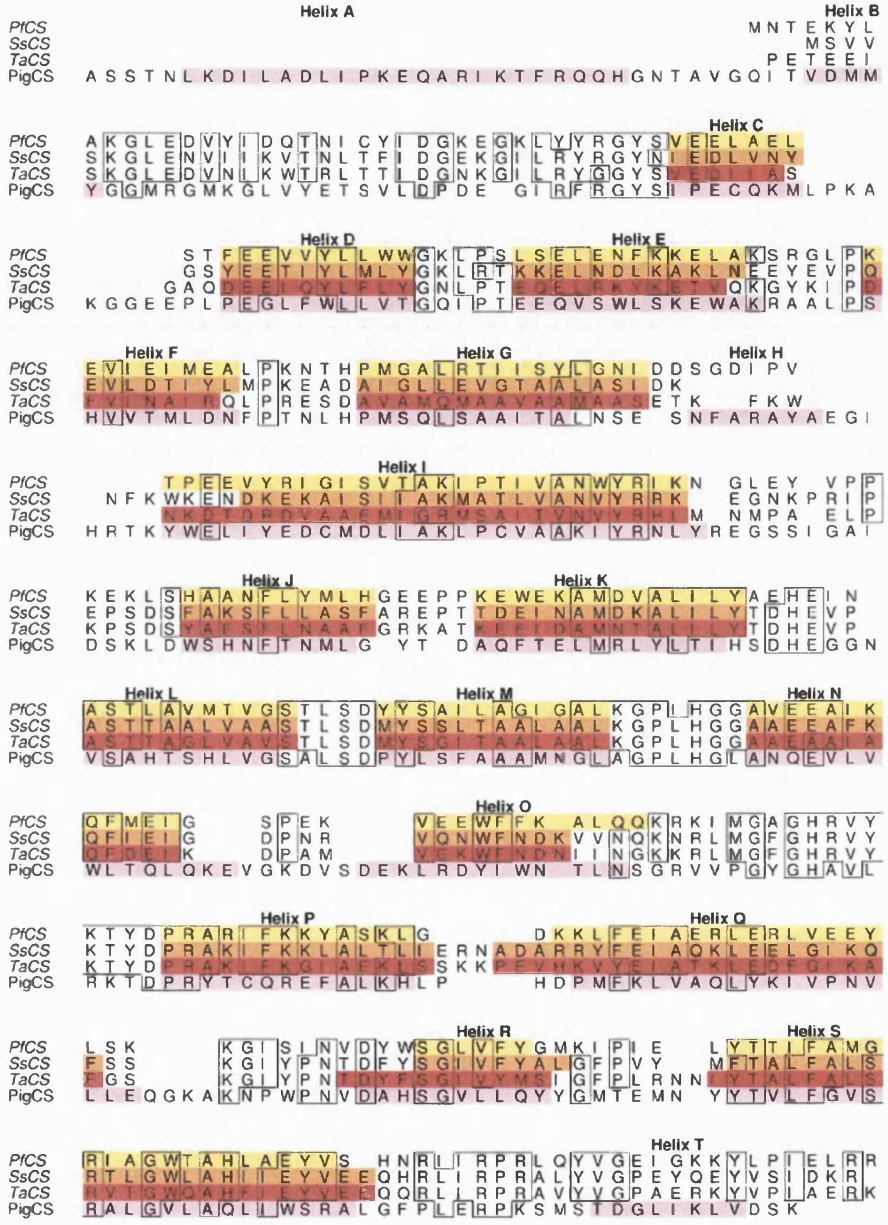


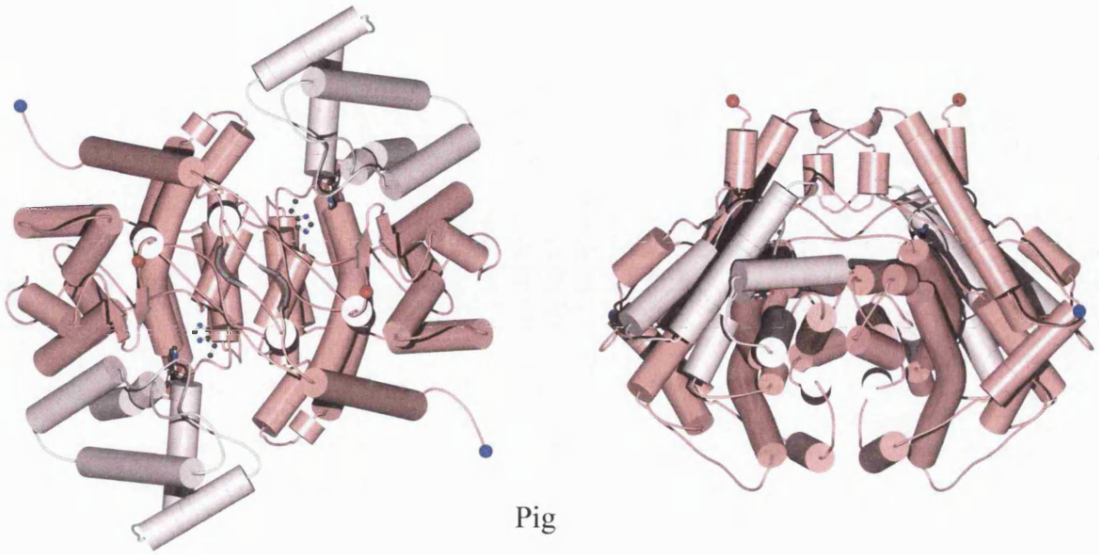
Figure 1.9 **Sequence comparison of pig and thermophilic citrate synthases**
Figure shows the structural alignment of the protein sequences from pigCS (helices highlighted in pink), *Ta*CS (red), *Ss*CS (orange), and *Pf*CS (yellow). The helices are labelled and highlighted.

The three-dimensional structure of these citrate synthase enzymes has been determined; *Thermoplasma acidophilum*, (open conformation/ligand-free) [Russell *et al.*, 1994], *Sulfolobus solfataricus,* (open conformation/ligand-free) [Bell, 1999], and *Pyrococcus furiosus*, (closed conformation/ligand-bound) [Russell *et al.*, 1997]. The crystal structures of these thermostable enzymes can be compared with mesophilic citrate synthase structures e.g. pig heart (37°C) (open and closed conformations) [Remington., 1982], to identify structural features of thermostability. All the CS structure data are summarised in Table 1.1. Figures 1.10(a-d) illustrate the high degree of structural homology of these citrate synthases.

Source	Opt. Growth Temp (°C)	Temp at which $t_{1/2}$=8min (°C)	Ligands	Data Res. (Å)	Reference
Pig (pigCS)	37	58	Citrate only Citrate and CoA	2.7 2.0	Remington *et al.*, 1982
T.acidophilum (*Ta*CS)	55	87	-	2.5	Russell *et al.*, 1994
S.solfataricus (*Ss*CS)	87	95	-	2.7	Bell, 1999
P.furiosus (*Pf*CS)	100	100	Citrate and CoA	1.9	Russell *et al.*, 1997

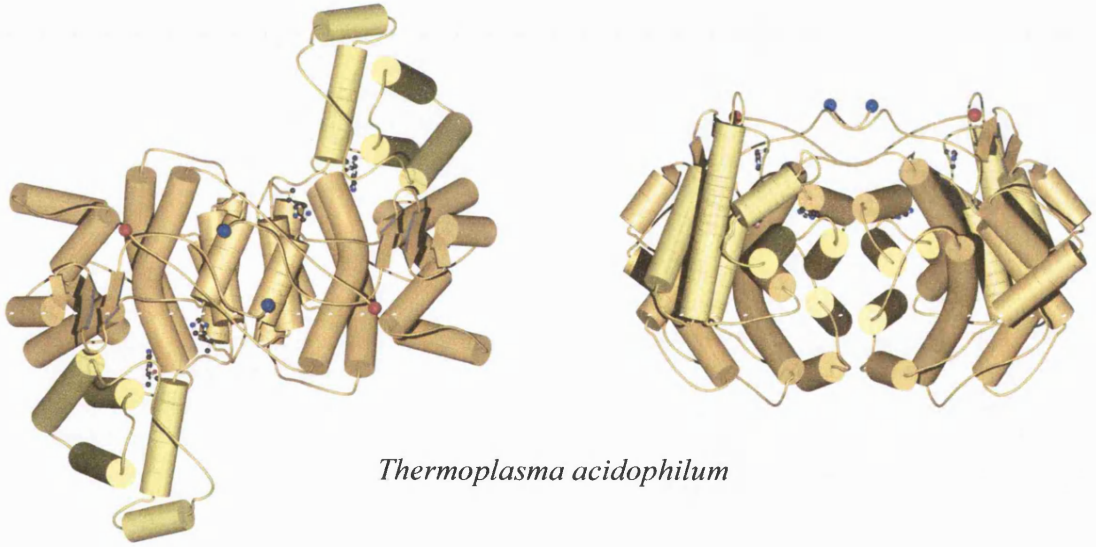
Table 1.1 Summary of citrate synthase structures used for structural comparisons.

(a)



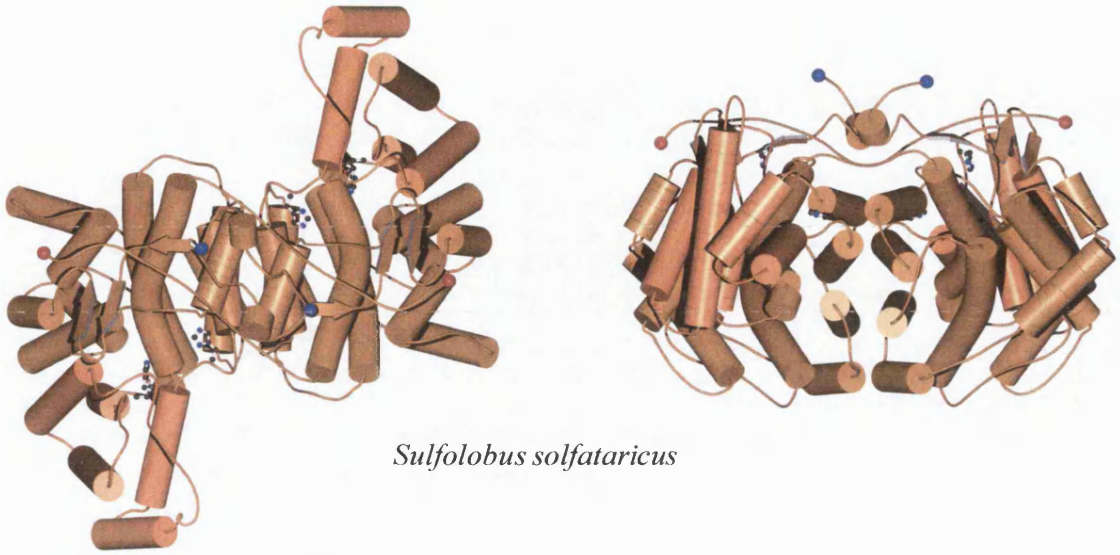
Pig

(b)



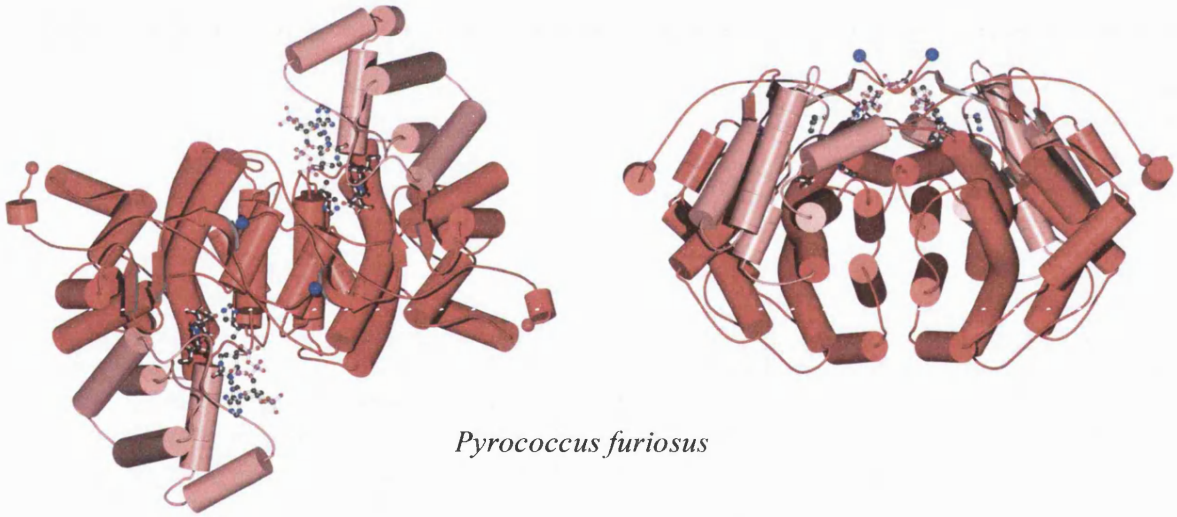
Thermoplasma acidophilum

(c)



Sulfolobus solfataricus

(d)



Pyrococcus furiosus

Figure 1.10 **Schematic representations of the mesophilic and thermophilic citrate synthase structures.**
In these figures the α-helices are shown as rods, and the catalytic residues and substrates are shown in ball-and-stick representation.

The dimer interface of all the CSs is made up of two parts: the main part being the eight α-helical sandwich of four antiparallel pairs of helices (F, G, L and M), with the additional interaction at the C-terminal regions. These features are discussed in further detail in Section 1.12.4, and in the introduction to Chapter 3. Figure 1.11 shows the three-dimensional structure of *Pf*CS, with the inter-subunit regions indicated. It is important to note that the C-terminal arms of the *Ta*CS and *Ss*CS are not complete in the three-dimensional models, as the last 14 and the last 8 residues respectively were not seen in the electron density map [Russell *et al.*, 1994; Bell, 1999]. This is probably due to the fact that the *Pf*CS was crystallized in the presence of ligands (closed conformation) whereas both the *Ta*CS and the *Ss*CS were crystallised in the open, ligand-free conformation. Hence there may be interactions present which have not been observed, also suggesting that the C-terminal arm seems to be ordered only in the presence of substrates.

Comparative structural analyses of these citrate synthase structures has revealed many features that may contribute to the thermal stability of this enzyme. Russell *et al.* (1994) compared the structures of citrate synthase from *Ta*CS and pigCS, and found that, despite the lack of sequence identity, the *Ta*CS had all the equivalent structural counterparts as the pigCS. Comparison of the *Ta*CS and pigCS structures with *Ss*CS [Bell,1999] and *Pf*CS [Russell *et al.*, 1997] identified further structural adaptations to stability at higher temperatures. The next section discusses the structural trends identified with increasing thermostability with reference to the citrate synthase enzyme [Bell, 1999; Arnott, 1999].

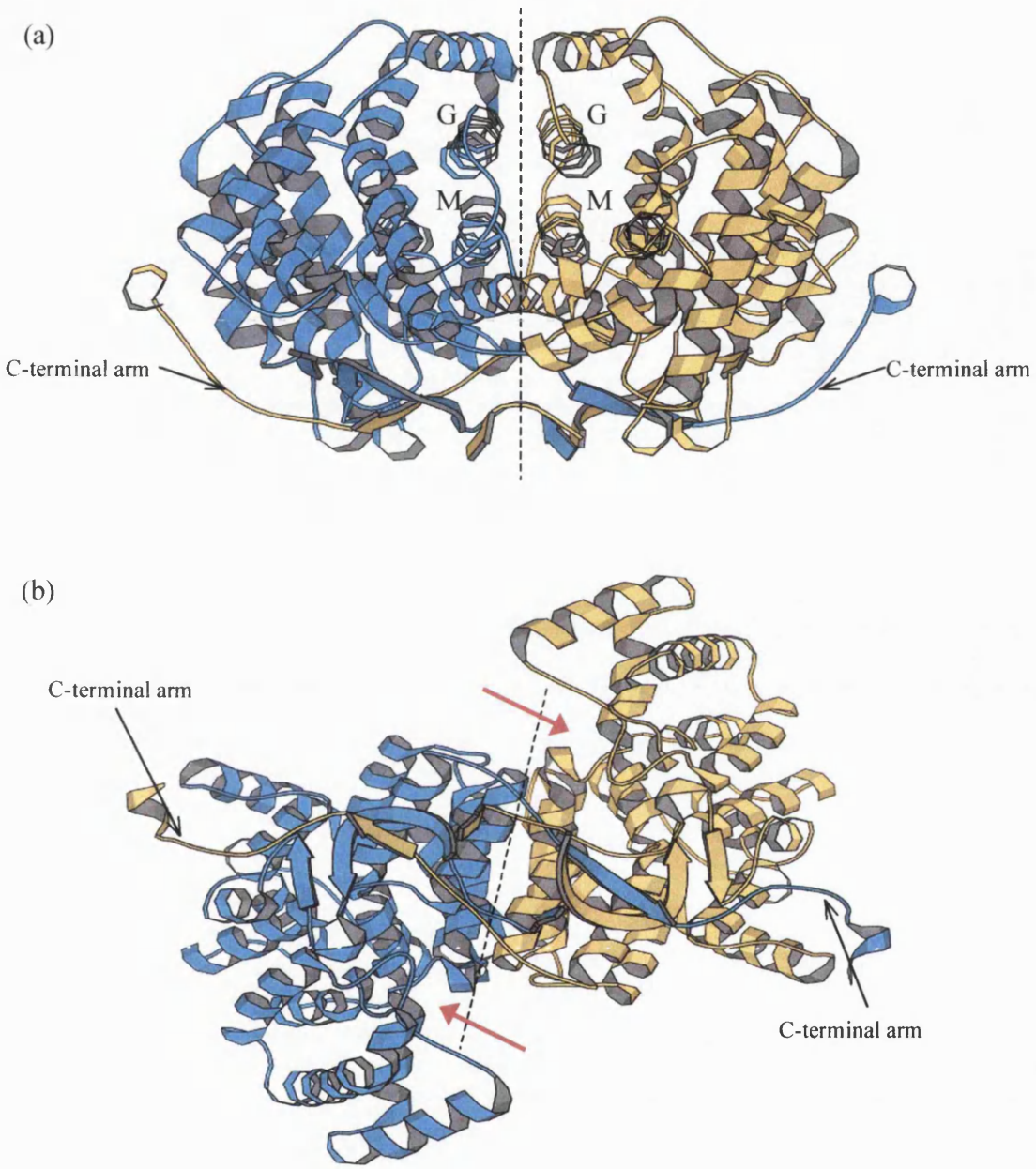


Figure 1.11 **Structure of *P.furiosus* citrate synthase.**

3.1(a) shows the crystal structure of *P.furiosus* citrate synthase, highlighting the subunit interface and C-terminal regions. Each monomer is shown in blue or yellow demonstrating how the C-terminal arm regions interact with the opposite monomer. The central helices (G and M) that contain the ionic networks are labelled. 3.1(b) is the 3.1(a) structure rotated –90° on its y-axis so the C-terminal arms can be illustrated clearly. The red arrows indicate the positions of the active sites, which lie between the large and small domain of each monomer.

1.12 General features identified from citrate synthase structural comparisons

1.12.1 Amino acid composition

In general the more thermostable citrate synthases have fewer thermolabile residues. These trends have been noted on comparing genome sequences of hyperthermophiles and mesophiles, as discussed in Section 1.6.1. In general the thermophilic structures have a higher content of charged residues (although lower content of histidine which is relatively thermolabile) than that of the mesophile. Many of the thermolabile residues found in pigCS are replaced in *Pf*CS by a residue that forms part of an ion pair, whilst some are eliminated via loop shortening. Of the hydrophobic residues, isoleucine seems to be favoured in the thermophiles but the overall hydrophobic content shows little difference. These substitutions result in the formation of isoleucine clusters sandwiched between the intersubunit helices (F, G, L and M). This clustering of isoleucine residues allows for a very well packed hydrophobic core. A similar clustering is observed in glutamate dehydrogenase from *P.furiosus* and has been implicated in the stability of this enzyme [Yip *et al.*, 1995]. Analysis of glycine and proline residues shows that there is no significant trend in the total contents of each amino acid, although location in the structure may be more important; the thermophilic citrate synthases have a lower number of intra-helical proline residues than pig CS.

1.12.2 Increased compactness

All thermostable citrate synthases have a considerably reduced accessible surface area and volume compared to the pigCS. There is also a decrease in the number of internal cavities present in the enzymes with increasing stability: 7 cavities in pigCS, 3 cavities in *Ta*CS, and no cavities present in *Ss*CS or *Pf*CS. Thus, the absence of large cavities in *Ss*CS and *Pf*CS may contribute to their thermostability by eliminating areas in the structure where non-optimal packing occurs (see Section 1.7.4). The presence of isoleucine clusters in the core of the *Pf*CS structure, as mentioned previously, may also contribute to optimising the packing of this enzyme.

Figure 1.12 shows that there is an overall shortening of loop regions in the CS structures with increasing thermostability. Compared to the pigCS structure, four loops have been identified in the *Ta*CS that are significantly shorter; the same four loops are also shortened in the *Pf*CS structure, and in addition a further two loops are also shortened compared to the pigCS.

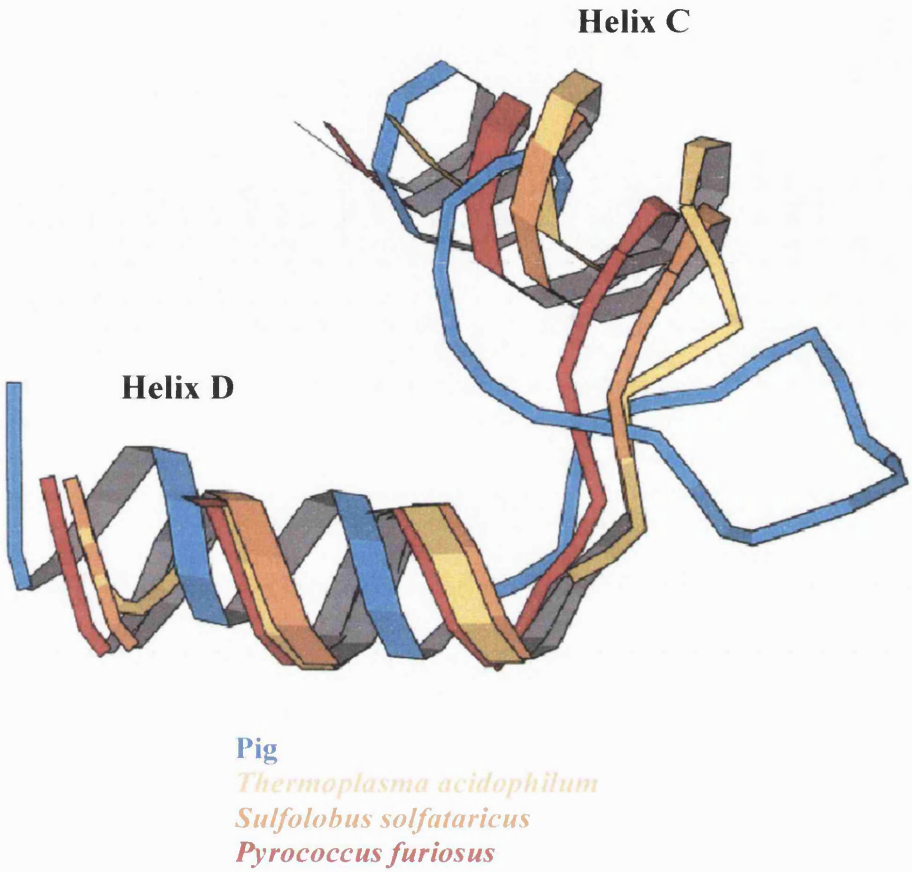


Figure 1.12 **Comparison of the loop regions connecting Helix C and Helix D over-laid for the citrate synthase enzymes.**

1.12.3 Increased number of ionic interactions

When the total number of ion pairs was calculated for pigCS, *Ta*CS and *Pf*CS using cut-off distances between oppositely-charged residues of 4, 5, and 6Å, there was observed to be a significantly increased number of total ion pairs found in *Pf*CS. A number of these ion pairs in *Pf*CS are located in loop regions, and the hypothesis is that the ionic interactions can 'pin' the loops down to the main structure to minimise flexible regions, as discussed in Section 1.7.1.

Looking simply at total number of ion pairs, it can be seen that, in general, all the thermophilic CSs have a greater total number of ion pairs than the pigCS, with *Pf*CS having the greatest number of inter-subunit ion pairs. The increase of ion pairs that can form ionic networks at the subunit interface regions of hyperthermophilic enzymes has been discussed previously (Section 1.7.3) and is discussed further in Section 1.12.4.

1.12.4 Improved homodimer association

The complementarity of packing at the subunit interface can be viewed by way of a gap volume index, defined as the interface gap volume divided by the interface surface accessible area. The gap volume index for pigCS is 1.74Å, while that for *Ta*CS is 1.03Å, *Ss*CS 1.26Å, and for *Pf*CS is 1.29Å, a smaller figure being indicative of closer packing [Bell, 1999]. However, in the three-dimensional structure of the open form of *Ta*CS the final 14 C-terminal residues are unresolved (possibly due to flexibility). Hence if the gap

volume index is recalculated for a 14 C-terminal deletion of *Pf*CS, then a figure of 1.02Å is obtained, which is consistent with the hypothesis of improved dimer contact with increasing thermostability. The gap volume index for the citrate synthase from the psychrophilic bacterium, DS23R, is 1.52Å which illustrates that the reduction of this value when comparing the thermostable enzymes with the mesophilic pig enzyme is not a result of comparing prokaryotic and eukaryotic enzymes.

Using a cut-off distance of 4Å, at each end of the central four helices, there are no ion pairs at the subunit interface of pigCS , but there is one ion pair at the subunit interface of *Ta*CS (although the interactions in this region are mainly hydrophobic), two ion pairs in *Ss*CS (although there are also many hydrophobic interactions) and a five-membered ionic network at the subunit interface of *Pf*CS (no hydrophobic interactions) (Figure 1.13). This is a common feature of hyperthermophilic enzymes, as discussed in Sections 1.7.3.

As discussed previously, in the *Pf*CS the monomer-monomer association is made up of two parts, the ionic network present between the 8 α-helical sandwich, and the additional interaction of the C-terminus of one monomer with the partner subunit. This C-terminal region consists of 13 residues and the penultimate residue Arg375 forms an ionic interaction with the Glu48 residue on the opposing monomer. The structure of the pigCS C-terminal region is different, see Chapter 3, and this region in the open crystal forms of the *Ta*CS and *Ss*CS is unresolved (see Section 1.11), possibly due to the flexibility of this region in the absence of substrates: again, these features are discussed in detail in Chapter 3.

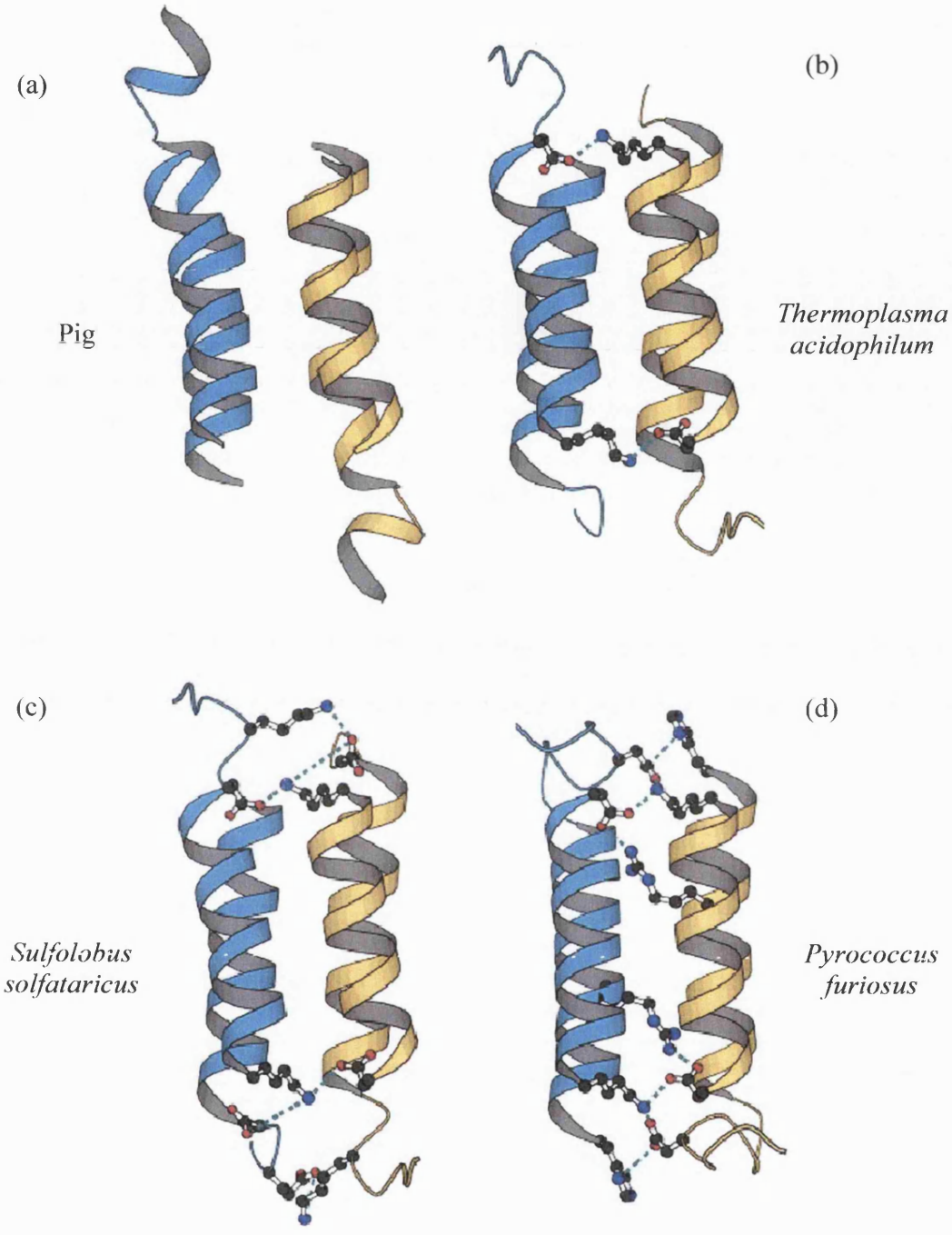


Figure 1.13 **Comparison of the central four helices of the citrate synthase enzyme.**
Figures show the 4 central helices, G, G', M, and M', of the inter-subunit region of citrate synthase. Helices from one subunit are shown in blue, the helices from the opposing subunit are in yellow

1.13 Experimental aims

Detailed analysis of the three-dimensional structures of citrate synthases, ranging from mesophilic to hyperthermophilic, has allowed the identification of structural trends with increasing thermostability. The structural trends identified in Section 1.12 are consistent with general trends of thermostability identified in other proteins, Section 1.7.

For this project, I am particularly interested in the investigation of the structural basis of hyperthermostability in citrate synthase from the hyperthermophilic *Pyrococcus furiosus*, and as discussed previously, an interesting feature of hyperthermostable enzymes is the increased number of ion pairs, particularly those that form ionic networks at the intersubunit regions.

The dimer integrity is particularly important in hyperthermostable citrate synthase as the active site comprises residues from both subunits, so the dimer must be maintained for catalytic activity. It is then unsurprising that the hyperthermophilic citrate synthase has evolved mechanisms to strengthen the interaction of the subunits.

This thesis describes and discusses the study of the inter-subunit ionic network and the C-terminal interactions of the *Pyrococcus furiosus* citrate synthase by site-directed mutagenesis, with two major aims:

1. To disrupt the C-terminal and ionic network regions in *Pf*CS to observe the effect on thermostability.

2. To introduce the five-membered ionic network present in *Pf*CS into the equivalent position in the less thermostable *Ta*CS to study the effects on thermostability.

CHAPTER 2

Materials and Methods

This chapter details the general techniques used in the experiments described in this thesis. Methods that are specific to a particular chapter are described within that chapter.

2.1.1 Reagents, enzymes and other materials

Molecular Biology: Vent DNA polymerase, the Rapid Ligase Kit (T4 DNA ligase), and Shrimp alkaline phosphate (SAP) were supplied by Boerhringer Mannhein, (Mannheim, Germany). Restriction enzymes were supplied by New England Biolabs, (Boston, USA). QUIAEX II Gel extraction Kit was supplied by Quiagen Ltd., (West Sussex, UK) and the Nucleospin Plus Plasmid Miniprep kit was supplied by Clonetech, (California, USA). Deoxynucleotide set was supplied by the Sigma Aldrich Chemical Company, (Poole, UK).

Cell culture: Bacto-agar, Yeast extract, Tryptone, Nalidixic acid, streptomycin sulphate, and ampicillin were all supplied by the Sigma Aldrich Chemical Company, (Poole, UK).

Electrophoresis: Agarose was supplied by Helena Biosciences, (Tyne and Wear, UK). The ethidium bromide and TEMED were supplied by the Sigma Aldrich Chemical Company, (Poole, UK). 1Kb DNA Ladder was supplied by GibcoBRL Life Technologies (Paisely, UK), and broad and long range protein markers were supplied by Bio-rad (California,

USA). Protogel™ acrylamide stock was supplied by National Diagnostics, (Georgia, USA). BioDesign GelWrap™ was supplied by Biodesign Inc., (New York, USA).

CS purification and assay: Dye Matrex Red Gel A was supplied by Millipore UK Ltd., (Watford, UK). Coomassie Brilliant Blue R, Coenzyme A (lithium salt from yeast), 5,5'-dithio-bis(2-nitrobenzoic acid) (DTNB), oxaloacetate, and acetic anhydride were supplied by the Sigma Aldrich Chemical Company, (Poole, UK). The Bio-rad Protein assay reagent was supplied by Bio-rad, (California, USA) and the Albumin Standard stock solution supplied by Pierce, (Illinois, USA).

All other chemical used were supplied by Fisons, (Loughborough, UK), Sigma Aldrich Chemical Company, (Poole, UK), Fluka Chemicals, (Gillingham, UK), and Merck, (Lutterworth, UK).

2.1.2 Bacterial strains, culture conditions and plasmids

The *E.coli* strain MOB154 (gltA6, galK30, pyrD36, relA1, rpsL129, thi-1 supE44, hsdR4, recA λ-) was grown in LB broth (1.0% (w/v) Bacto-trypton, 0.5% (w/v) yeast extract, and 1.0% (w/v) NaCl) at 37°C, shaking at 200rpm, or on LB plates, 1.5% Bacto agar, or in Terrific Broth (1.2% (w/v) Bacto-trypton, 2.4% (w/v) yeast extract, 0.4% (v/v) glycerol, 17mM KH_2PO_4, and 72mM K_2HPO_4). For the growth of *E.coli* strains carrying plasmids conferring ampicillin resistance, media was supplemented with ampicillin to give a final concentration of 100μg/ml. Bacterial strains expressing streptomycin resistance genes were

grown on media and on plates supplemented with 25μg/ml streptomycin sulphate. All bacterial strains were stored as glyceral stocks (15% v/v) at –70°C.

The pREC7/*NdeI* vector was kindly supplied by Dr Linda C. Kurtz of Washington University School of Medicine, (St Louis, USA).

2.2 Molecular Biology Methods

2.2.1 Polymerase chain reaction (PCR)

Reactions were carried out in 0.5ml Eppendorf tubes, which were placed in an Eppendorf Mastercycler. The PCR reaction mix contained 10pmoles of a forward and reverse primer, 1 unit of Vent DNA polymerase, ×10 Vent DNA polymerase buffer (supplied with the Vent DNA polymerase), the template DNA to be amplified (0.1μg) and 50μM of each dNTP, unless otherwise stated. Each reaction was made up to a final volume of 100μl with sterile distilled water. The PCR machine was programmed to heat for 5 min at 96°C, followed by a step of 3 min at 85°C for the addition of the Vent DNA polymerase (hot start). These steps were followed by 30 cycles of the following conditions;

96°C for 1.25 min (denaturing)

55°C for 2 min (annealing)

72°C for 2 min (extension)

72°C for 10 min (final step after 30 cycles)

This was followed by a cooling step of 4°C until the samples were retrieved for analysis. Where the PCR conditions were different, details are given in the relevent chapters. Samples of the PCR products were taken, and 1 × DNA loading buffer was added (see next section), and the products run on a 1% (w/v) agarose gel.

2.2.2 Agarose gel electrophoresis

Horizontal agarose gel elctrophoresis was used for purification and analysis of small DNA fragments. Agarose was dissolved in 1 × TAE buffer (40mM Tris-acetate, pH 8.0, 1mM EDTA) by heating in a microwave oven. When the agarose had cooled to approximately 45°C, ethidium bromide was added to a concentration of 0.5μg/ml. The agarose was poured into a gel mould containing a well-forming comb, and the gel left to set at room temperature. Once set, the comb was removed and the gel placed in a gel tank and covered with 1 × TAE buffer. DNA samples in 1 × DNA loading buffer (0.04% (w/v) bromophenol blue, 0.04% (w/v) xylene cyanol, 5% (v/v) glycerol), were loaded into the wells and electrophoresed at a constant voltage of 80-100V, alongside a 1Kb ladder DNA marker, until the xylene cyanol dye front was 1-2cm from the edge of the gel. The DNA separated was visualised under UV illumination. DNA bands were cut out of the agarose gel for purification (see following section).

2.2.3 DNA purification from agarose gel

Following horizontal agarose gel electrophoresis, the relevant band was visualised using a UV transilluminator and excised from the gel using a scalpel. The DNA was then purified from the gel using the QUIAEX II Gel extraction kit. The excised agarose gel was dissolved in QXI solubilising and binding buffer, 3 × QXI to 1 × volume of gel. QUIAEX II DNA binding silica particles were added depending on the amount of DNA to extract (10μl for <100ng, 30μl for >100ng of DNA). This mixture was incubated at 50°C for 10 min. The mixture was spun at 13,000 rpm in a microfuge for 30 secs and the supernatent removed. The pellet is then washed with 500μl of QXI, followed by another spin. This is followed by two rounds of the pellet being resuspended in an ethanol buffer then spun in the microfuge. The pellet was air-dried for 10-15 min, then the DNA eluted by resuspending the pellet in 20μl od sterile, distilled water and spinning at 13,000 rpm for 30 secs.

2.2.4 Restriction Digestion of DNA

Restriction digestion was carried out in accordance with the enzyme manufacturer's instructions. The incubation time and the number of units of enzyme used were varied depending on the amount of DNA being cut, usually 1 unit of enzyme/1μg of DNA. Digests for the identification of successfully ligated and tranformed plasmid DNA were incubated at 37°C for 3 hrs. Double digests for the ligation of inserts and vector were usually incubated at 37°C overnight. All citrate synthase genes were cloned into the

pREC7/*NdeI* vector using the restriction enzymes, *NdeI* and *KpnI*. When double digests were carried out, the NE Buffer 1 restriction buffer, 10mM Bis Tris propane-HCl, 10mM $MgCl_2$, 1mM dithiothreitol (pH 7.0 at 25°C), that was compatible with both *NdeI* and *KpnI* was used. The NE Buffer 1 was also supplemented with 100μg/ml BSA to aid the enzymic activity of the *KpnI* enzyme. Under these conditions the *KpnI* restriction enzyme displays 100% activity, whereas the *NdeI* restriction enzyme displays 75% activity.

2.2.5 5' Phosphorylation of oligonucleotides

50-100ng of oligonucleotide were phosphorylated , prior to ligation, by the addition of 5 units of Shrimp alkaline phosphatase (SAP), ×10 dephosphorylation buffer (supplied with SAP), to the restriction digest mixture. Each reaction was incubated at room temperature for 10-30 min, and was stopped by heating at 65°C for 15 min. Phosphorylated oligonucleotide was stored at –20°C.

2.2.6 DNA ligation

Ligations were carried out in a final volume of 20μl using the Rapid DNA ligase Kit by Boerhringer Mannheim, (Mannheim, Germany). The recommended insert to vector ratio was 1:3 vector : insert (50ng of linearised phosphorylated vector and 150ng of insert DNA). The ligations were incubated at room temperature for 30 min. Each reaction contained 5 units of T4 DNA ligase and 10μl of 2 × T4 DNA ligation buffer in a total volume of 21μl.

Controls: 1) no ligase and no insert (to check for uncut vector DNA), 2) no insert (checks for vector religation).

2.2.7 Transformation of competent cells

MOB154 *E.coli* cells were made competent and transformed with double stranded DNA using the calcium chloride method. A 1ml LB overnight culture of MOB154 cells was used to innoculate a 50ml LB culture. The cells were incubated at 37°C, shaking at ≥200rpm, until a final OD_{650nm} equalled 0.4. The cell culture was chilled on ice for 5 min, then spun at 4000g for 10 min at 4°C. The harvested cells were resuspended in 20ml of ice-cold 50mM $CaCl_2$, and left on ice for 2 min. This was followed by another spin at 4000g for 10 min at 4°C, then the pellet was finally resuspended in 5ml of 50mM $CaCl_2$ supplemented with 15% (v/v) glycerol. This suspension was aliquotted into 100µl aliquots and snap-freezed to be stored at –70°C.

100µl of competent cells were used in each transformation reaction. 50-100ng of DNA, either from a ligation reaction (5-10µl) or uncut, was added to the competent cells and left on ice for a minimum of 30 min, before heat shock treatment in a water bath at 45°C for 45 secs. The cells were aliquotted into 0.8ml of LB and incubated at 37°C at 200rpm in an incubator for 90 min before plating onto LB plates containing the appropriate supplements. The plates were incubated overnight at 37°C and transformed colonies selected.

2.2.8 Extraction of plasmid DNA from cell cultures

Plasmid DNA was extracted from 1.5ml LB overnight cultures of transformed colonies using the Nucleospin Plus Plasmid Miniprep Kit supplied by Clonetech, (California, USA). The 1.5ml LB overnight culture is spun at 3000-5000g for 5-10 min, then the pellet resuspended in 250µl of Buffer A1 (resuspension buffer). 250µl of Buffer A2 (lysis buffer) was added and incubated at room temperature for 5 min. To remove protein from the cell extract, 300µl of Buffer A3 (binding buffer) was added and incubated on ice for 5 min. The cell extracts were then spun at maximum speed in a microfuge for 10 min at 4°C, and the supernatant applied to a Nucleospin column. The column was washed several times with an ethanol buffer (70% ethanol), and the DNA was finally eluted from the column with 50µl of distilled water. Samples were stored at –20°C.

2.2.9 DNA sequencing

DNA was sequenced using the Perkin Elmer ABI 377 Auotomated DNA sequencer. Samples were prepared for sequencing by aliquotting 50-200ng of 'miniprep' DNA, and 3-10pmoles of oligonucleotide primer into an 0.5ml microfuge tube. The volume was made up to 6µl using sterile distilled water. Samples were elongated using T7 DNA polymerase and fluorescent labels. All citrate synthase genes were sequenced using both forward and reverse primers. The Wisconsin GCG Sequence Analysis software package was used analyse sequencing data.

2.3 Protein expression and purification

2.3.1 Protein Expression

For large scale protein purification (30ml Dye Matrex Red A column), the bacteria were grown overnight in 50ml of Terrific Broth containing 100μg ampicillin/ml. 1L of Terrific Broth (100μg ampicillin/ml) was innoculated with the 50ml overnight culture and grown to a OD_{600} of 5, when the cell culture was induced with a final concentration of 50μg nalidixic acid/ml (to induce the pREC7/*NdeI* vector). The cell culture was incubated at 37°C for a further 16-18 hrs, and the cells were harvested by spinning at 3000-5000g for 30 min at 4°C. For small-scale purification (2ml Dye Matrex Red A column), 1ml of overnight culture was used to innoculate 50ml Terrific Broth, 100μg ampicillin/ml.

2.3.2 Preparation of cell extracts

Harvested cells were resuspended in 20mM EPPS, pH 8.0, 2mM EDTA, 20ml per 2-3g of cells. The resuspension was sonicated for 15-20 sec bursts 3 times, with incubation on ice between bursts, then centrifuged at 12,000g for 15 min.

2.3.3 Protein purification

The prepared cell extracts were subjected to a heat step; 15 min at 85°C for *Pf*CS wt, 80°C for the *Pf*CS mutants, and 65°C for all the *Ta*CSs, followed by centrifugation at 13000 × g for 20 min. Each extract was then loaded either onto a 2ml (small-scale), or 30 ml (large-

scale) Dye Matrex Red A column, equilibrated with 20mM EPPS, pH 8.0, 2mM EDTA (loading buffer). In smallscale citrate synthase purifications unbound protein was removed by washing the column with 10 × column volume of loading buffer. In large-scale purifications, unbound protein was removed by washing the column with loading buffer until the A_{260nm} baseline returned to zero.

CS was eluted with loading buffer containing 5mM oxaloacetate and 1mM CoA [James *et al.*, 1994]. 1ml fractions were collected and assayed for CS activity, and samples were run on 10% (w/v) SDS-PAGE to check for purity. The pure CS tractions were pooled.

2.4 Biochemical techniques

2.4.1 Determination of protein concentration

Protein concentration was determined by the method of Bradford (1976), using the Bio-rad protein assay stock reagent diluted 1:4 reagent:distilled water, then filtered. A 0.9ml volume of the diluted, filtered Bio-rad protein assay reagent was incubated with a 0.1ml volume of protein samples at room temperature for 15 min. The absorbance at 595nm was measured and the protein concentration calculated from a calibration curve prepared with Albumin standard solutions (0-10μg/ml)

2.4.2 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Protein samples were analysed by the SDS-PAGE method. A 20ml resolving gel solution (10% (w/v) Protogel™ acrylamide, 0.375M Tris-HCl pH8.9, 0.1% (w/v) SDS) was polymerised by the addition of 5mg of ammonium persulphate (from a 10% (w/v) solution), and 12.5µl of TEMED, and poured into a gel mould. The gel solution was overlaid with water-saturated n-butanol and left to polymerise at room temperature. The n-butanol layer was removed and the top of the gel washed with distilled water. A 10ml stacking gel solution (5% (w/v) Protogel™ acrylamide, 0.13M Tris-HCl, pH 6.5, 0.1% (w/v) SDS) was polymerised with 3mg ammonium persulphate (from a 10% (w/v) solution) and 10µl of TEMED. The stacking gel was poured to the top of the mould, and a comb placed into the gel, and allowed to set for 1 hour at room temperture. The comb was then removed carefully, and the gel placed in an electrophoresis tank, which was filled with an electrophoresis buffer (0.025M Tris-HCl, pH 8.8, 0.096M glycine, 0.1% (w/v) SDS). Unpolymerised acrylamide was washed from the wells of the stacking gel with electrophoresis buffer prior to loading of the protein samples.

Protein samples were prepared for electrophoresis by the addition of an equal volume of SDS-PAGE loading buffer (0.18M Tris-HCl, pH 6.5, 2% (w/v) SDS, 10% (w/v) glycerol, 2.6% (w/v) β-mercaptoethanol, 0.005% (w/v) bromophenol blue), and heated at 95°C for 5 min. 10µg of cell extracts and 4µg of purified protein were run on the gel. The gel was run at a constant current of 10mA through the stacking gel and 20mA through the resolving gel until the dye front was 0.5cm from the lower edge of the gel. The gel was then removed

from the electrophoresis tank and stained for 10 min with 0.5% (w/v) Coomassie Brilliant Blue R in 9:2:9 (v:v:v) methanol:acetic acid:water and then destained overnight with 2:1:7 (v:v:v) methanol:acetic acid:water.

2.4.3 Spectrophotometric assay for citrate synthase

Citrate synthase was assayed at 412nm in a Perkin Elmer Lambda 11 spectrophotometer, using the method of Srere *et al.* (1963). The thiol group of the reaction product (free CoA) reacts with 5,5'-di-thiobis-2-nitrobenzoic acid (DTNB) to release thionitrobenzoate which absorbs at 412nm and has an absorption coefficient of $13,600M^{-1}cm^{-1}$.

Coenzyme A was acetylated by dissolving 10mg of coenzyme A in 1ml of Milli-Q water (on ice), and adding $KHCO_3$ to a concentration of 0.17M to bring the concentration to approximately pH 8.0 from pH 5.0, and acetic anhydride to a concentration of 11mM. The solution was incubated on ice for 20 min to allow complete acetylation of the coenzyme A.

Unless otherwise stated, the enzyme was assayed in 50mM EPPS buffer, (pH 8.0), 2mM EDTA and 100mM KCl, with 0.2mM oxaloacetate and 0.14mM acetyl-CoA.

2.4.3(i) Determination of kinetic parameters

For the determination of kinetic parameters, assays were carried out in varying concentrations of one substrate whilst the second substrate was fixed at a high

concentration (~10-20 × K_m). K_m and V_{max} values were calculated from these data by the direct linear plot of Eisenthal & Cornish-Bowden (1974) using Enzpack 3 computer software (Biosoft, Cambridge UK).

2.4.3(ii) Thermal inactivation

For the *Pf*CS wt and mutants, thermal inactivation studies were carried out by incubating aliquots of citrate synthase (100μg/ml) in 50 mM sodium phosphate buffer, pH 7.0, 2mM EDTA, 0.1M KCl, in sealed glass capillary tubes. Tubes were incubated at the defined temperature in a PEG bath and removed at known intervals and rapidly cooled in an ice/water bath. For the *Ta*CS wt and mutants, thermal inactivation studies were carried out by incubating 100μl aliquots of citrate synthase (~30μg/ml) in 50 mM sodium phosphate buffer, pH 7.0, 2mM EDTA, 0.1M KCl in 0.5ml PCR tubes. Tubes were incubated at the defined temperature in the Eppendorf Mastercycler and removed at known intervals and rapidly cooled in an ice/bath. Enzymic activity was then measured under normal assay conditions at pH 7.0.

2.4.3(iii) Temperature optimum

The dependence of catalytic activity on temperature was studied by assaying citrate synthase at temperatures between 40 and 90°C. The *Pf*CSs were assayed in 50mM sodium phosphate buffer, pH 7.0, 2mM EDTA, 0.1M KCl, and the *Ta*CSs were assayed in 50mM EPPS buffer, pH 8.0, 2mM EDTA, 0.1M KCl. The concentrations of acetyl-CoA and

OAA were chosen to remain at saturating concentrations throughout the assay, and are defined in the text where relevant.

CHAPTER 3

Investigating the role of the inter-subunit ionic network and the C-terminal region in the thermal stability of *Pyrococcus furiosus* citrate synthase

3.1 Introduction

Pyrococcus furiosus is a hyperthermophilic organism isolated from marine sediment surrounding geothermal sulphurous volcanoes. It is an anaerobic organism that grows optimally at 70-103°C and pH 5-9 [Fiala *et al.*, 1986]. The gene encoding citrate synthase from *P.furiosus* has been cloned and overexpressed in *E.coli* and the recombinant enzyme purified [Muir *et al.*, 1995, James *et al*, 1994]. The ligand-bound (closed) crystal structure has been determined to a resolution of 1.9Å [Russell *et al.*, 1997]. The enzyme consists of two identical monomers of 42.6 kDa [Muir *et a*l., 1995]. As previously discussed, each monomer is composed of 16 α-helices distributed between a large and small domain, and the interface of the two subunits is composed of an 8 α-helical sandwich of 4 antiparallel pairs of helices (F, G, M and L).

As discussed previously, several trends were observed when comparing the crystal structures of citrate synthase from *P.furiosus*, *S.solfataricus, T.acidophilum*, the mesophilic pig and the psycrophilic Athrobacter DS23R. Generally, there is an overall increased compactness of the enzyme, a reduction in loop regions and a reduction in the number of

thermolabile residues. More significantly there is a more intimate association of the subunits including an increase in the number of intersubunit ion pairs.

The major structural trend of interest identified concerns the nature and strength of interactions at the subunit interface. Preliminary studies of citrate synthase using differential scanning calorimetry (DSC) suggest that the thermal inactivation of the enzyme may involve the dissociation of the two subunits with subsequent unfolding of the individual monomers. The citrate synthase active site requires residues from both monomers; hence the maintenance of dimer integrity at high temperatures is essential.

This chapter primarily studies the role of features of the subunit interface interactions as a structural basis of thermostability, and two aspects have been investigated. Firstly, compared to the less thermostable citrate synthases, the number of possible ion pairs is increased in the *Pyrococcus furiosus* citrate synthase (*Pf*CS) enzyme. These additional ion pairs are located primarily in Helix G and Helix M that make up the central four helices of the eight helical sandwich at the subunit interface (Figure 3.1), and importantly interact with each other to form a network.

Yip *et al.* (1998) used a homology-based modeling study to compare sequences of a range of glucose dehydrogenase enzymes (GDHs) from species spanning the biological temperature range. They analysed the extent of ion-pair network formation and correlated the data with observed thermal stability, showing that the networks become more fragmented as the temperature optimum of the enzyme decreases.

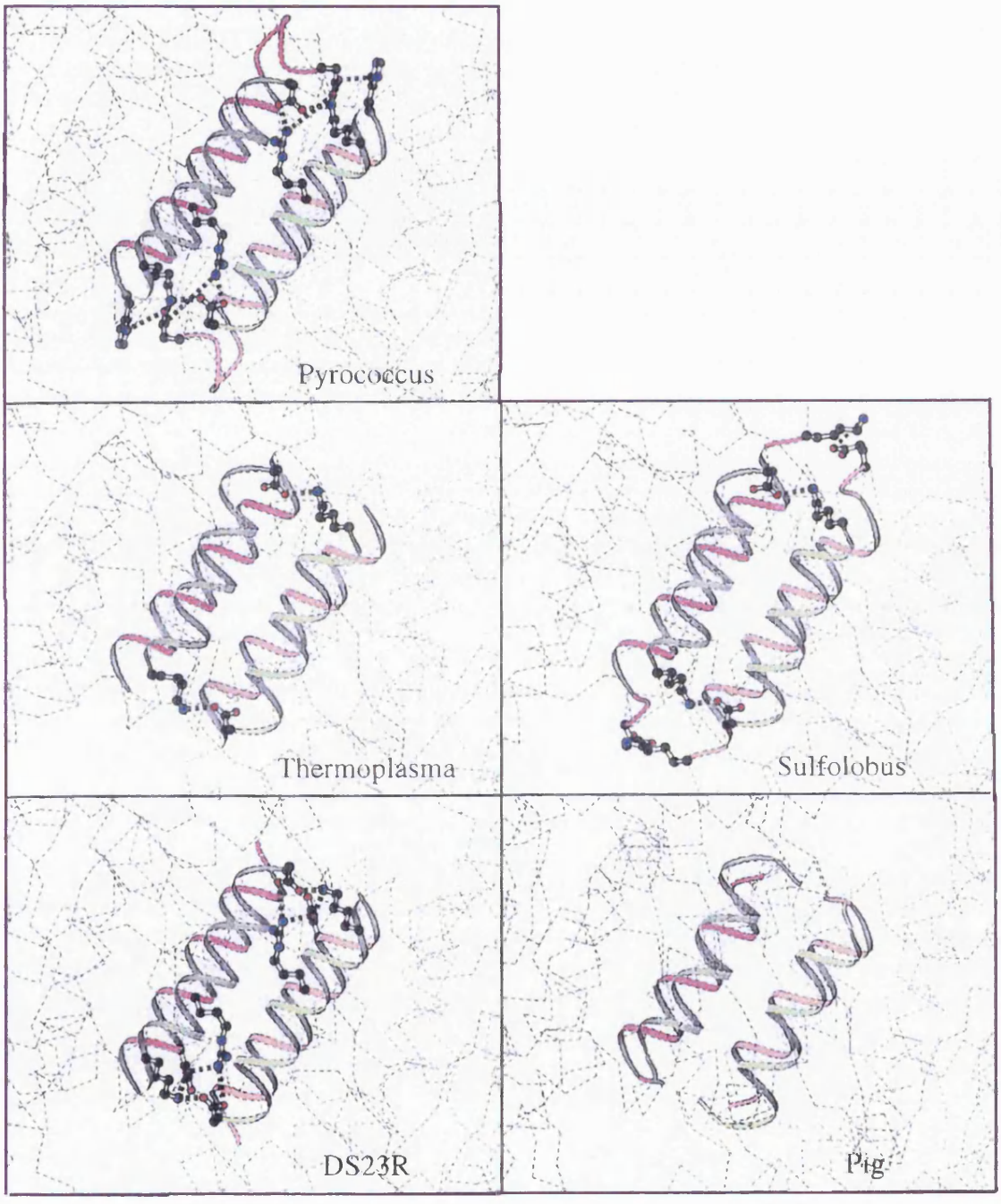


Figure 3.1 **Comparison of the ionic network regions of several citrate synthases.**

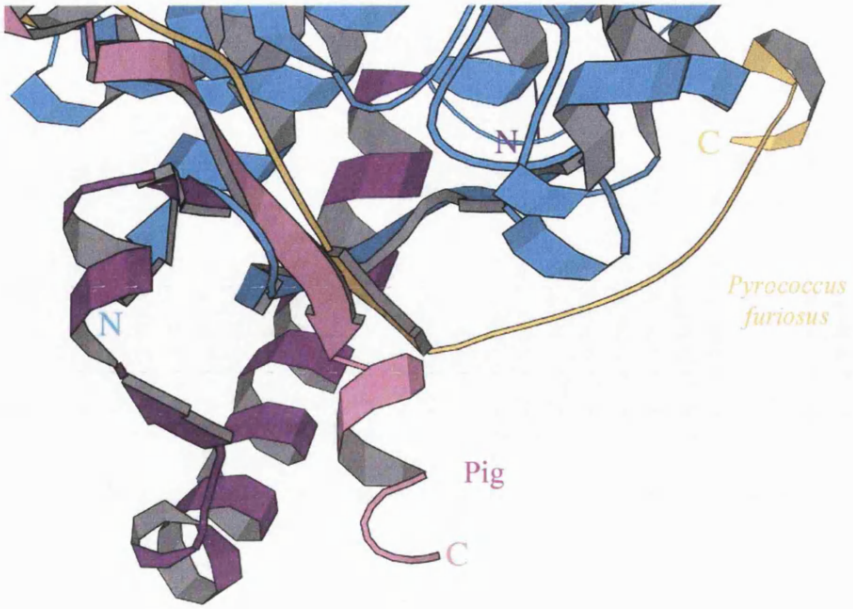
Diagram showing the ionic interactions in the central helices (G (green) and M (brown)) of the dimer interface.

This is consistent with the theory that such networks play an important role in the adaptation of enzymes to extreme temperatures.

Ionic interactions have been extensively studied in the glutamate dehydrogenase enzyme (GDH) [Aghjanian *et al.*, 1998]. Removal and substitution of ionic residues in this enzyme have showed that ionic networks of the mutant have a significant effect on the thermal stability of GDHs compared to the wild-type enzyme [Rahman *et al.,* 1998, Vetriani *et al.*, 1998].

Secondly, the crystal structure of the closed form of *Pf*CS wt [Russell *et al.*, 1997] has revealed that the C-terminal arm of one monomer folds in an extended form over the surface of the other subunit to give the appearance of the monomers embracing one another. Figure 3.2(b) shows that the *Pf*CS has two terminal arginine residues. Examination of the crystal structure revealed that the penultimate residue (Arg375) formed an ionic interaction with a glutamate residue (Glu48) on the other associated monomer. It was predicted that this interaction might contribute to maintaining the integrity of the dimer interface at high temperatures. In comparison with less thermostable citrate synthases it can be seen that this region varies greatly in structure (Figure 3.2). This C-terminal region is shorter in the pigCS by 4 residues (Figure 3.2(b)) but, as shown in Figure 3.2(a), this region forms a small helix, which is completely different to the *Pf*CS C-terminal arm.

(a)



(b)

DS2-3R	L	S	E	Y	N	G	P	E	Q	R	Q	V	P						
Pig	K	S	M	S	T	D	G	L	I	K	L	V	D	S	K				
Thermoplasma	R	A	V	Y	V	G	P	A	E	R	K	Y	V	P	I	A	E	R	K
Sulfolobus	R	A	L	Y	V	G	P	E	Y	Q	E	Y	V	S	I	D	K	R	
Pyrococcus	R	L	Q	Y	V	G	E	I	G	K	K	Y	L	P	I	E	L	**R**	**R**
						↑												↑	
						-13												-2	

Figure 3.2 **Comparison of C-terminal region of citrate synthases.**

3.3(a) is a ribbon diagram of the C-terminal regions of *P.furiosus* and Pig citrate synthase. The *Pf*CS is shown with the A subunit in yellow and the B subunit in blue, and the pigCS C-terminus is shown in pink (this region packs against the N-terminal extension of the other subunit shown in purple). 3.3(b) is a sequence alignment of the C-terminal arms of various citrate synthases. The 13 amino acid arm of *Pf*CS is indicated as are the terminal arginine residues (bold).

Figure 3.2(a) does not include the C-terminal arms structures of *Ta*CS wt and *Ss*CS wt as these regions are unresolved in the open crystal structures, although the sequence alignment in Figure 3.2(b) shows high homology between the thermostable citrate synthases in this region. This may suggest that the C-terminal region only becomes ordered in the presence of substrates as the *Pf*CS was crystallised in the closed, ligand-bound form whereas the *Ta*CS and *Ss*CS were crystallized in the open, ligand-free form. This may suggest that, if the arm region is present in the *Ta*CS and *Ss*CS open structures, the absence of bound substrates may cause this region to be in a more flexible conformation. Preliminary data from the closed, ligand-bound crystal of *Ta*CS shows that the C-terminal arm region is partially more ordered in the presence of substrates (Rupert Russell, personal communication).

C-terminal regions have been shown to affect thermostability in other proteins. Hamana *et al.* (1998) analysed the contribution of C-terminal amino acids to the activity and thermostability of *T.thermophilus* orotate phosphoribosyl transferase, ORPTase, by constructing ORPTase mutants with amino acid deletions. Results showed that deletions in the C-terminal region lowered the thermostability of this enzyme.

This chapter investigates the role of the C-terminal arm region and the extensive ionic network in the thermostability of the *Pf*CS by site-directed mutagenesis.

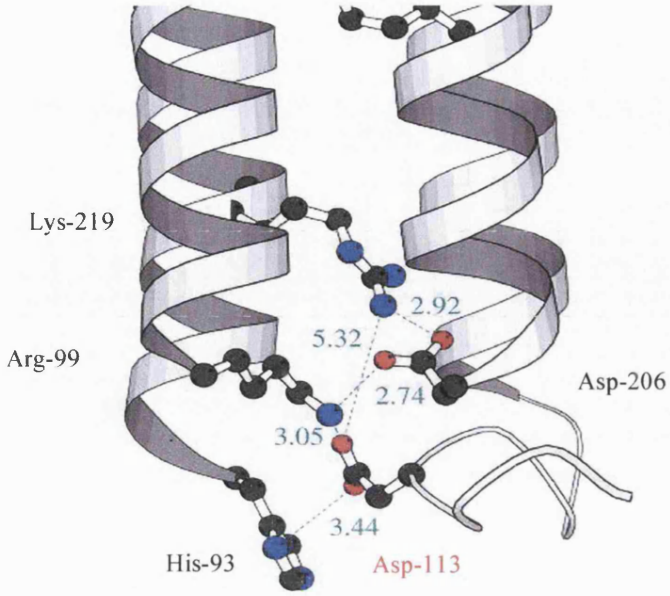
3.2 Previous work

3.1.1 *Pf*CS ionic network mutants.

Arnott (1999) constructed two *Pf*CS mutants to study the role of the ionic network in the hyperthermostability of *Pf*CS. As discussed previously (Section 3.1) evidence suggests that the ion-pair network present at the subunit interface of *Pf*CS may contribute to the increased thermostability of this enzyme.

Residues His93, Arg99, Asp113, Asp206 and Lys219 from both subunits form a five-membered ionic network at both ends of the central 4 helices (G, G', M and M') of the 8 α-helical sandwich of the subunit interface. Figure 3.3 illustrates the positioning of these residues and the helix and subunit each residue is on. Arnott (1999) constructed two mutant *Pf*CSs, both of which consisted of the replacement of Asp113 on Helix G. This residue was selected as it would disrupt 3 ionic interactions and it is also fairly exposed to solvent compared to the more buried members of the network (see the accessibility values on Figure 3.3(b)). This substituted aspartate residue is an acidic residue which carries a negative charge. The first mutation has the aspartate replaced with the polar serine residue, the larger, hydroxyl version of alanine; this mutant is termed *Pf*CS(D113S). The hypothesis of this substitution was that the serine might be able to retain a degree of interaction with the other residues in the ionic network, as it is a polar residue, so complete disruption of the ionic network may be avoided. The second mutation constructed was the

(a)



(b)

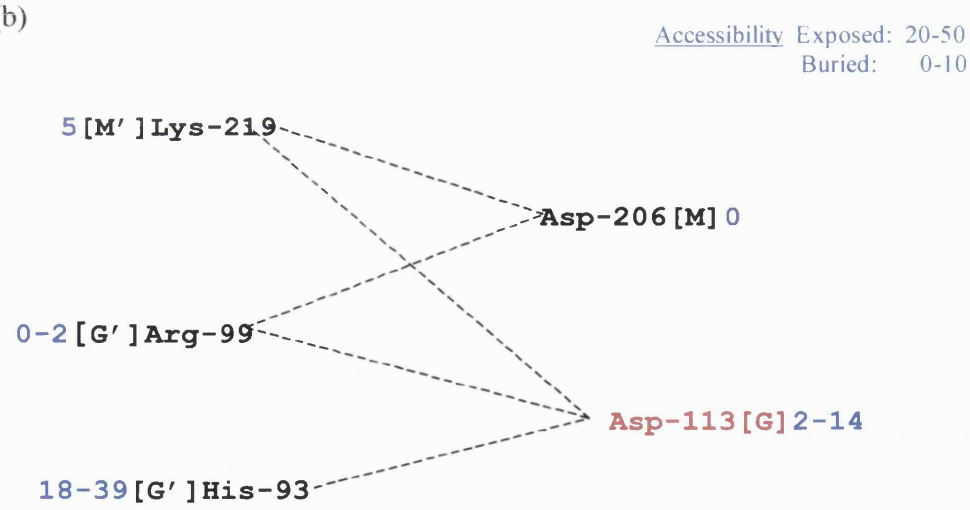


Figure 3.3 **Diagram of the ionic network at the subunit interface of *Pf*CS**
(a) shows the 3-dimensional positioning of the ionic network between helices M, M', G and G'. All ionic interactions are shown with bond lengths in Å. (b) is a schematic of the ionic network from an end on view of the helices indicating the helices on which each residue lies. The residue substituted is shown in red. The accessibility values (blue) indicate the exposure of the residue to solvent (using the programme NACCESS) where 0-10 is regarded as a buried residue and 20-50 as exposed.

replacement of Asp113 with a small, hydrophobic alanine residue; this mutant is termed *Pf*CS(D113A). This substitution was intended to disrupt all ionic interaction of the Asp113 as the alanine residue should be unable to form any ionic bonds or polar interactions.

The mutants were constructed by digesting the mutated region out of the gene products of the nested PCR and ligating them into a corresponding digest of the *Pf*CS wt gene. All *Pf*CSs were cloned into the pREC7/*NdeI* expression vector [Arnott, 1999].

3.2.2 *Pf*CS C-terminal deletion mutants.

The C-terminal arm region of *Pf*CS consists of 13 amino acids and forms an extended arm that wraps around the opposite monomer (Figure 3.2). As this was a potential contribution to the association of the subunits, C-terminal deletion mutants were constructed to test this hypothesis.

The terminal arginine residue (Arg375) forms an ionic interaction with a glutamate residue (Glu48) on the other associated monomer. Arnott (1999) constructed a *Pf*CS mutant [*Pf*CS (-2)] where these two terminal arginine residues were deleted to remove the interaction with the other monomer. A second *Pf*CS mutant [*Pf*CS (-13)] was constructed where the entire 13-residue C-terminal arm region was deleted to prevent any interaction between subunits in this region, *Pf*CS (-13).

These mutations were introduced by PCR using primers that incorporated stop codons upstream of the bases coding for the C-terminal regions. Again these *Pf*CS C-terminal deletion mutants were cloned into the pREC7/*NdeI* expression vector [Arnott, 1999]. All *Pf*CS gene/vector constructs were transformed into a citrate synthase negative *E.coli* strain, MOB154 for expression.

In earlier work by Arnott (1999) characterization of the *Pf*CS subunit interaction mutants was incomplete; hence, this chapter describes and discusses the kinetic and thermal characterization of these mutants.

Results

3.3 Transformation of *Pf*CS wt and the *Pf*CS (D113S) mutant into MOB154 cells.

The *Pf*CS gene and the *Pf*CS ionic network and C-terminal deletion mutants had previously been cloned into the pREC7/*NdeI* expression vector using *NdeI* and *KpnI* restriction sites (see Previous work, section 3.2). All of the *Pf*CS ionic network and C-terminal deletion mutants (apart from the *Pf*CS wt and the *Pf*CS (D113S) ionic network mutant) were available in MOB154 citrate synthase negative *E.coli* glycerol stocks. The *Pf*CS wt and *Pf*CS (D113S) mutant were only available as gene/vector constructs. Hence, both the *Pf*CS wt and the *Pf*CS (D113S) gene/vector constructs were transformed into MOB154 cells.

Five colonies of *Pf*CS wt transformed cells and six colonies of *Pf*CS (D113S) transformed cells were selected after incubation on LB-amp agar plates overnight. The plasmid DNA was isolated from cultures of each colony. Figure 3.4 shows the *NdeI* and *KpnI* restriction digests of the plasmid DNA from the five *Pf*CS wt selected colonies. The gene insert is ~1.3kb and the vector is 2.5kb, hence the final construct is 3.8kb. Figure 3.4 shows that all selected colonies have been successfully transformed with the correct construct, having bands corresponding to the insert, vector and the partially digested full construct. A digest of the plasmid DNA from the 6 selected *Pf*CS (D113S) transformed colonies showed exactly the same results as for the *Pf*CS colonies; hence this gel is not shown. For both constructs, one colony was selected for expression and purification.

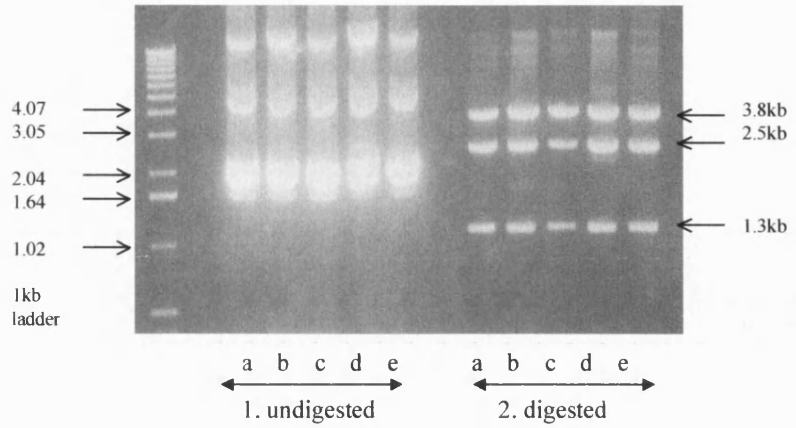


Figure 3.4 **Restriction digest of *Pf*CS wild-type in pREC7/*NdeI*.**

This figure shows the plasmid DNA prepared from 5 colonies of MOB154 cells transformed with the *Pf*CS wt gene/vector construct. Lanes 1(a-e) correspond to the undigested plasmid from the selected colonies a-e. Lanes 2(a-e) show the plasmid digested with *NdeI* and *KpnI* restriction enzymes. The 3.8kb band corresponds to the singly cut full construct, the 2.5kb band corresponds to the pREC7/*NdeI* expression vector, and the 1.3 kb band corresponds to the *Pf*CS gene insert. Colony (a) was selected for expression of *Pf*CS wt.

3.4 Expression and purification of *Pf*CS wt and the *Pf*CS inter-subunit contact mutants.

The transformed MOB154 cells for *Pf*CS wild-type and all the *Pf*CS mutants were grown in a 1L culture, under control of the recA promoter on the pREC7/*NdeI* vector, and induced with 50μg nalidixic acid/ml. In each case between 8-14g of cells were obtained after centrifugation. Heat-treated cell-free extracts were prepared and applied to a 30 ml Dye Matrex Red A column. The citrate synthase was eluted with one bed volume of 5mM OAA and 1mM CoA. Each 1ml fraction collected was assayed for citrate synthase activity and the active fractions were pooled together. Table 3.1 shows the activity and protein concentration of all *Pf*CS at each stage of purification.

Enzyme	Stage of purification	Volume (ml)	Total activity (U)	Total protein (mg)	Specific activity (U/mg)	Yield (%)	Purification fold
*Pf*CS wt	Cell extract	39.0	530.3	285.1	1.9	100.0	-
	Heat treated	36.0	402.7	117.4	3.4	75.9	1.8
	Red A extract	18.0	156.5	16.9	**9.3**	**29.5**	**5.0**
*Pf*CS (–2)	Cell extract	39.0	9852.5	496.1	19.9	100.0	-
	Heat treated	37.0	5935.2	204.2	29.1	60.3	1.46
	Red A extract	25.0	4669.9	145.3	**32.1**	**47.4**	**1.62**
*Pf*CS(–13)	Cell extract	37.0	199.2	280.5	0.7	100.0	-
	Heat treated	23.0	64.4	24.4	2.6	32.3	3.7
	Red A extract	18.0	110.8	20.0	**5.5**	**55.6**	**7.8**
*Pf*CS (D113A)	Cell extract	40.0	1640.6	252.4	6.5	100.0	-
	Heat treated	38.0	1405.3	66.1	21.3	85.7	3.3
	Red A extract	9.0	537.6	14.0	**38.4**	**32.8**	**5.9**
*Pf*CS (D113S)	Cell extract	40.0	1072.7	627.3	1.7	100.0	-
	Heat treated	37.0	641.4	88.1	7.3	59.8	4.3
	Red A extract	12.5	117.1	8.1	**14.5**	**10.9**	**8.5**

Table 3.1 **Purification table of *Pf*CS wild-type and *Pf*CS subunit interface mutants.**
Table shows the total activity (U) and the total protein (mg) present at each stage of the purification process; crude cell extract, heat treatment at 85°C (for wild-type) or 80°C (for mutants), and purification on a Dye Matrex Red A column. Table indicates the yield of citrate synthase obtained at each stage. Citrate synthase activity is assayed as standard (55°C, pH 8.0) and total activity is quoted in Units (μmoles product/min).

All of the citrate synthases are expressed as soluble, active proteins. Figure 3.5(a) shows an SDS-PAGE of the cell extract, heat-treated extract, and purified citrate synthase collected from the Dye Matrex Red A column for wild-type *Pf*CS. Figure 3.5(b) shows that all the purified enzymes were homogenous on SDS-PAGE and the enzymes had Mr values of ~43 kDa [Muir *et al.*, 1995].

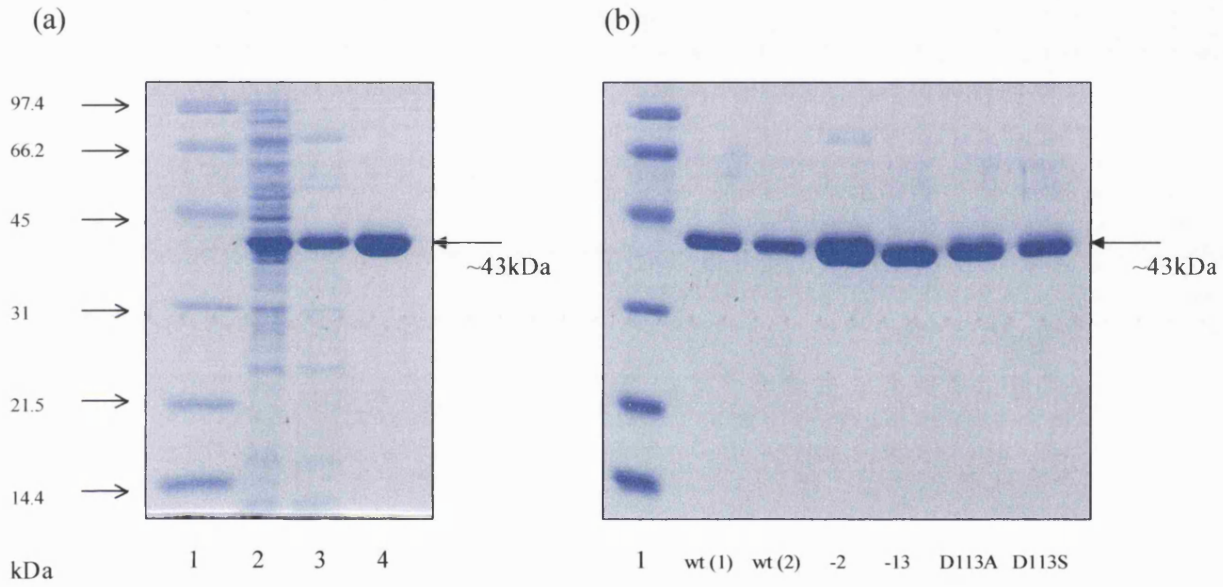


Figure 3.5 **SDS-PAGE of purification of *Pf*CS wild-type and mutants.**
Figure (a) shows the purification of *Pf*CS wild-type. Lane 1 contains protein markers (long Mr range). Lane 2 corresponds to the cell-free crude extract, lane 3 corresponds to the cell extract after heat treatment for 15 min at 85°C, and lane 4 corresponds to the *Pf*CS after purification on the Matrix Red Gel A column. Figure (b) shows the *Pf*CS wild-type (expressed in pKK223 (1) and pREC7/*NdeI* (2) expression vectors) and all of the *Pf*CS mutants after treatment on the Dye Matrex Red A column to illustrate similar degrees of purity.

3.5 Kinetic parameters of *Pf*CS wt and the *Pf*CS inter-subunit contact mutants

Preliminary data on the kinetic parameters of the *Pf*CS wt and each of the *Pf*CS mutants had previously been obtained by Arnott (1999). A more detailed analysis was required, therefore, for each citrate synthase; the K_m for oxaloacetate and acetyl-CoA and the V_{max} of the enzyme were determined. The K_m for each enzyme was determined by adding a fixed excess of one substrate (20 × K_m, based on the preliminary results previously obtained) and varying the concentration of the other (0.5, 0.7, 1, 5 and 10 × K_m). The data were analysed using ENZPACK, and the K_m and V_{max} data were obtained using the Direct Linear plot [Eisenthal & Cornish-Bowden, 1974] with 68% confidence limits. Figure 3.6 shows the data for *Pf*CS wildtype as a Michaelis-Menten plot for both substrates, OAA and AcCoA. Figure 3.7 shows the same data as a Hanes-Wolfe plot (S/V vs S), as a straight-line plot can best judge the linearity of the data.

The V_{max} values were calculated using data obtained from the ENZPACK Direct Linear plot and data from Bradford protein estimations of the enzyme samples. The errors for the V_{max} values were determined by combining the errors from the ENZPACK data and the errors from the protein estimation (determined in triplicate for each sample), by using the equation:

$$\frac{(\Delta Z)^2}{(Z)^2} = \frac{(\Delta A)^2}{(A)^2} + \frac{(\Delta B)^2}{(B)^2}$$

Where Z = V_{max} (μmol/min/mg), A = enzyme velocity (μmol/min/ml) and B = protein concentration (mg). The Δ symbol indicates the error (or change) for each value.

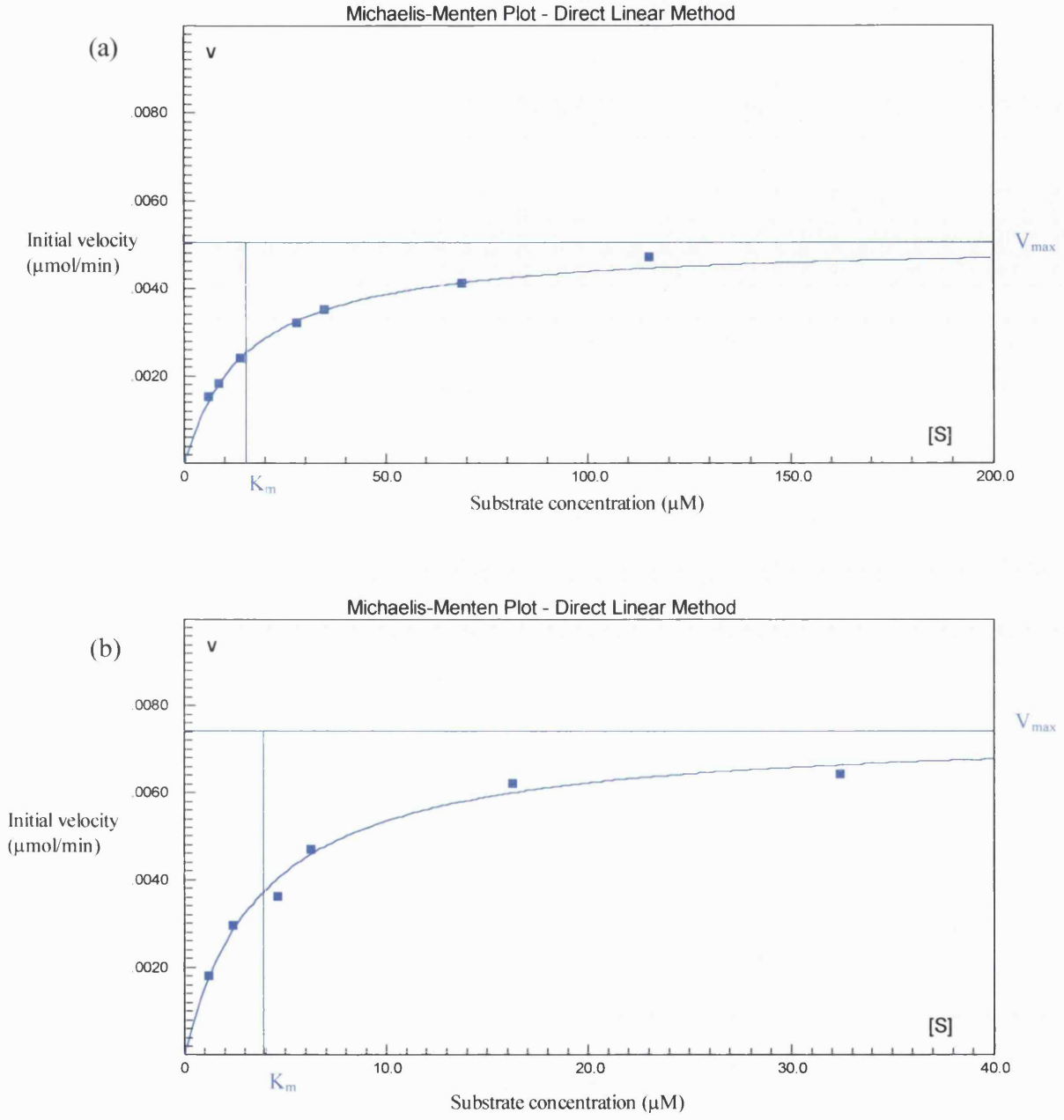


Figure 3.6 **Michaelis-Menten plots of *Pf*Cs wt.**
Figure (a) shows the velocity vs [S] plot of *Pf*CS wt when AcCoA is fixed (>10×K_m), 0.06mM, and the concentration of OAA is varied to give to K_m value for OAA. Figure (b) shows the velocity vs [S] plot for *Pf*CS wt when OAA is fixed (>10×K_m), 0.2mM, and the concentration of AcCoA is varied to give the K_m value for AcCoA. Line is generated, in ENZPACK, from K_m and V_{max} values determined by the Direct Linear Plot. Assays contained a final enzyme concentration of ~0.2µg/ml.

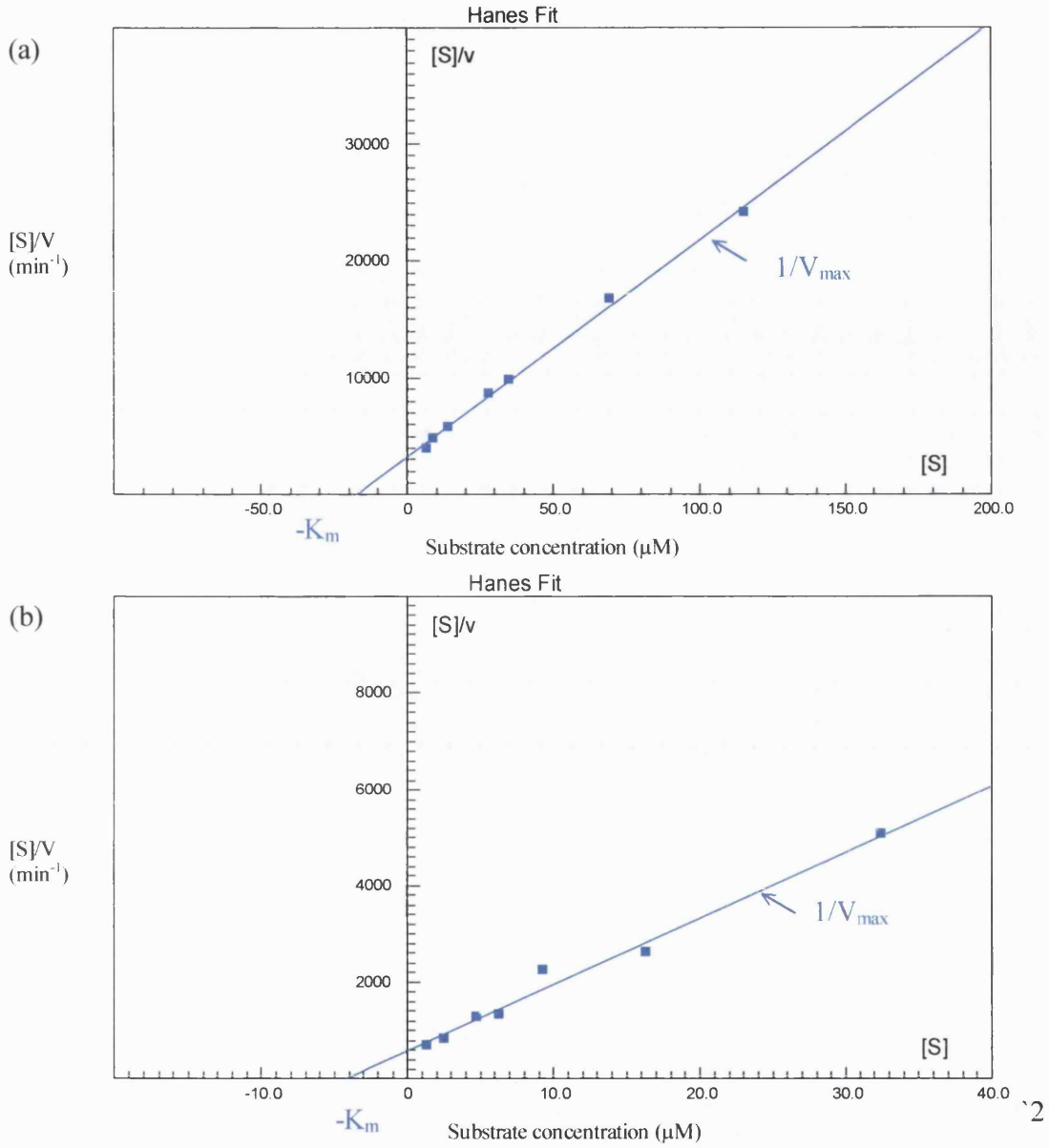


Figure 3.7 **Hanes plot of PfCS wt**

Figure (a) shows the [S]/V vs S plot of *Pf*CS wt when AcCoA is fixed (>10 $\times K_m$), 0.06mM, and the concentration of OAA is varied to give a K_m value for OAA. Figure (b) shows the [S]/V vs S plot for *Pf*CS wt when OAA is fixed (>10 $\times K_m$), 0.2mM, and the concentration of AcCoA is varied to give the K_m value for AcCoA. Plots generated in ENZPACK. Assays contained a final enzyme concentration of ~0.2μg/ml.

Table 3.2 summarises the kinetic data for *Pf*CS wt and the *Pf*CS ionic network and C-terminal deletion mutants.

Enzyme	Km AcCoA (μM)	Km OAA (μM)	Vmax (μmol/min/mg)	k_{cat} (s^{-1})
*Pf*CS wild-type	4 (±0.3)	15 (±1)	23 (±1)	16
*Pf*CS (-2)	5 (±0.5)	20 (±1)	17 (±1)	12
*Pf*CS (-13)	51 (±5.0)	212 (±18)	14 (±1)	10
*Pf*CS (D113A)	7 (±0.3)	17 (±1)	24 (±1)	17
*Pf*CS (D113S)	7 (±0.6)	12 (±1)	19 (±1)	14

Table 3.2 **Kinetic parameters determined for *Pf*CS wt and *Pf*CS mutants.**
Table shows the K_m and V_{max} values determined for *Pf*CS wild-type and the 4 subunit interaction mutants. The K_m values are determined by fixing the concentration of one substrate (>20×K_m) and varying the concentration of the other. The V_{max} values were calculated using data from the Direct Linear plot and determination of protein concentration (mg). The standard errors quoted for the K_m and V_{max} values are calculated from the 68% confidence limits from the Direct Linear plot. The k_{cat} values are calculated using a Mr value of 42.6 in each case.

As the active sites of *Pf*CS lie at the subunit interface any structural changes induced in the mutant enzymes (all mutations are in this subunit interface region) may affect the kinetic parameters of the enzyme. However, Table 3.2 shows that the ionic network mutants (*Pf*CS (D113A) and *Pf*CS (D113S)) have kinetic constants that are similar to that of the wild-type enzyme. This is also true in the case of one of the C-terminal deletion mutants *Pf*CS (-2). However *Pf*CS (-13) possesses K_m values that are increased compared to the *Pf*CS wild-type, although it has retained V_{max} and k_{cat} values comparable to *Pf*CS wt. Given the general sensitivity of kinetic parameters to structural changes at the active site, it

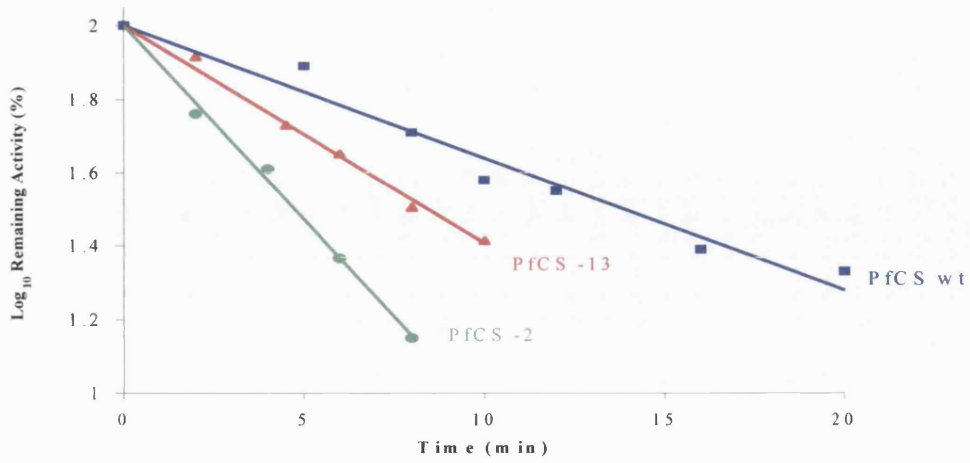
is felt that the data in Table 1 imply that the mutations have not induced large conformational change in the enzyme, except perhaps in the case of *Pf*CS (-13).

3.6 Thermal inactivation of *Pf*CS wt.

The *Pf*CS wild-type and the *Pf*CS mutants were analysed for their thermostability properties in terms of their resistance to irreversible heat denaturation. Each enzyme was incubated at ~100°C (100°C for all thermal inactivations apart from the C-terminal mutants in the absence of substrates, which were incubated at 103°C) and, at defined time intervals, samples were snap-cooled in an ice-water bath and then assayed for enzyme activity at 55°C under standard assay conditions (pH 7).

Previous thermal inactivation data were obtained by Arnott (1999) where the citrate synthase enzymes were incubated at ~100°C in the absence of substrates (Figure 3.8). Data of this nature are more difficult to compare directly with temperature (T_{opt}) experiments or kinetic analyses that are carried out in the presence of both citrate synthase substrates (OAA and AcCoA). Therefore, the thermal inactivation experiments were repeated under the same conditions but this time with the addition of OAA and CoA during the incubation (Figure 3.9). An increased concentration of OAA was used to ensure that the enzymes were fully saturated despite the thermal instability of this substrate. Also CoA, one of the citrate synthase products, was used instead of AcCoA to prevent any reaction occurring during the incubation period.

(a)



(b)

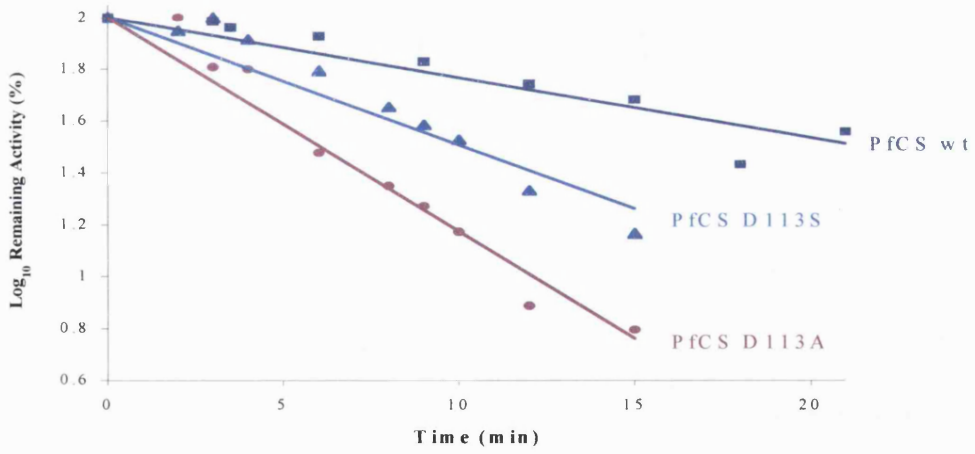
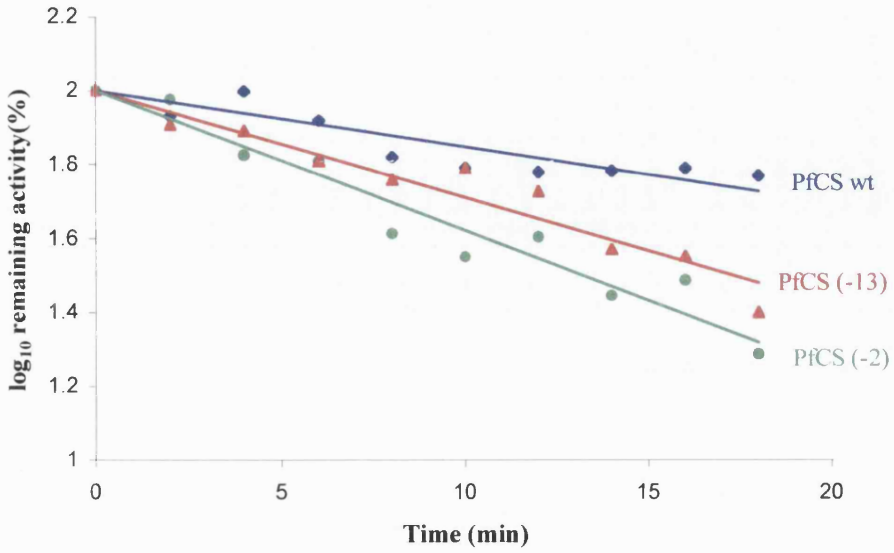


Figure 3.8 **Thermal inactivation of *Pf*CS wt and subunit interface mutants in the absence of substrates.**
Enzyme samples were incubated at a high temperature and the remaining activity was measured at known time intervals under standard assay conditions at 55°C. (a) shows the thermal inactivation of *Pf*CS wt and the C-terminal mutants at 103°C and (b) shows the thermal inactivation of *Pf*CS wt and the ionic network mutants at 100°C. Data produced by Arnott (1999).

(a)



(b)

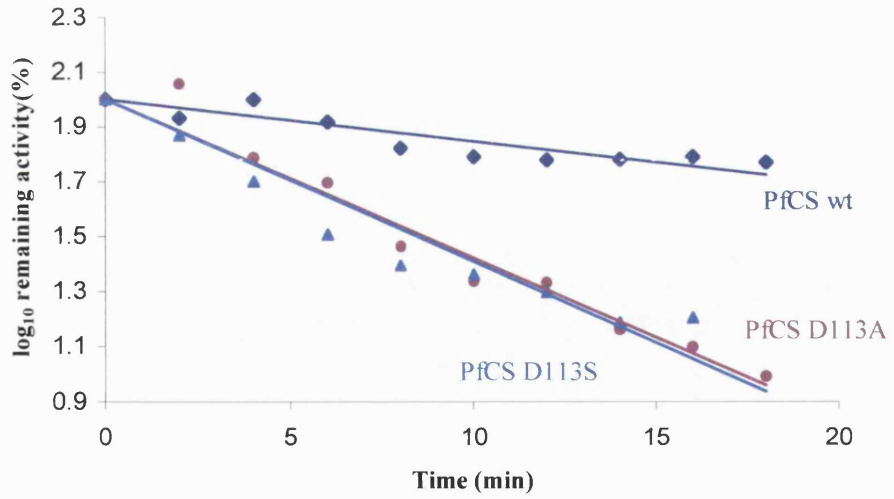


Figure 3.9 **Thermal inactivation of *Pf*CS wt and the subunit interface mutants in the presence of substrates.**
Enzyme samples were incubated at 100°C in the presence of 5mM OAA and 0.14mM CoA, and the remaining activity was measured at known time intervals under standard assay conditions at 55°C.

Table 3.3 shows the 1st order rate constants (k_{inact}) and the half-lives ($t_{½}$) for the thermal inactivations of *Pf*CS wild-type and the *Pf*CS mutants. All of the *Pf*CS mutants show an increase in the k_{inact} values over that of the *Pf*CS wt enzyme, indicating an increased rate of thermal inactivation at 100°C. These values correspond to a shorter $t_{½}$ at this temperature.

	Thermal inactivation in the absence of substrates		Thermal inactivation at 100°C in the presence of substrates	
Enzyme	k_{inact} (min^{-1})	$t_{½}$ (min)	k_{inact} (min^{-1})	$t_{½}$ (min)
*Pf*CS wt	0.083 (± 0.005)* 0.053 (± 0.005)†	8.4 13.1	0.035 (± 0.002)	20.1
*Pf*CS (-2)	0.244 (±0.002)*	2.8	0.088 (± 0.005)	7.9
*Pf*CS (-13)	0.136 (± 0.005)*	5.1	0.067 (± 0.002)	10.4
PfCS (D113A)	0.191 (± 0.018)†	3.6	0.134 (± 0.005)	5.2
*Pf*CS (D113S)	0.113 (± 0.005)†	6.1	0.134(± 0.007)	5.1

Table 3.3 **Thermal inactivation first order rate constants (k_{inact}) and half-lives ($t_{½}$) of *Pf*CS wt and subunit interface mutants.**

k_{inact} and $t_{½}$ data were calculated from the slopes of the thermal inactivations at 100°C as shown in Figures 3.8 and 3.9. Two values for *Pf*CS wt are shown, * indicates the values obtained from Figure 3.8(a) incubated at 103°C for the inactivation of the C-terminal mutants. † indicates the values obtained from the thermal inactivation at 100°C (Figure 3.8(b)) for the inactivation of the ionic network mutants. Enzyme samples were incubated at the relevant temperature in the absence and presence of 5mM OAA and 0.14mM CoA, and the remaining activity was measured at known time intervals under standard assay conditions at 55°C.

Figure 3.8(a) shows that both of the C-terminal deletion mutants have a greater rate of thermal inactivation than the *Pf*CS wt. Surprisingly, removal of the two terminal arginine residues has greatly decreased the thermostability of *Pf*CS wt (a 2.9-fold increase in k_{inact}). Some of this stability seems to be recovered with the removal of the entire 13-residue C-terminal region that shows a 1.6-fold decrease in k_{inact} compared with *Pf*CS (-2). These data

may be explained by assuming that the removal of the terminal arginines disrupts the interaction of Asp375 with Glu48 on the opposing subunit. This may cause an increase in the flexibility of the C-terminal arm region, which may be a target for thermal inactivation.

Disruption in the ionic network of *Pf*CS wt also results in an increase in the rate of thermal inactivation. Figure 3.8(b) shows that *Pf*CS (D113A) has an increased rate of thermal inactivation compared to both *Pf*CS wild-type and *Pf*CS (D113S), a 3.6-fold and 1.7-fold increase respectively. This is consistent with the hypothesis that the introduction of an alanine residue to replace Arg113 would disrupt all interactions with the rest of the ionic network. The *Pf*CS (D113S) is more stable at 100°C than *Pf*CS (D113A) but still less stable than the *Pf*CS wild-type. This may suggest that the serine substituted into the ionic network may retain some interaction with other charged residues, therefore retaining some thermal stability.

Interestingly the k_{inact} values of all the citrate synthases show a decrease in the presence of substrates, which suggests that the substrates are able to provide a degree of protection against thermal inactivation for these enzymes (Figure 3.9). This is most apparent in *Pf*CS (-2), which shows a 2.7-fold decrease in k_{inact} in the presence of substrates. This is important as Figure 3.9(b) shows that there is no significant difference in the thermal denaturation between *Pf*CS (D113A) and *Pf*CS (D113S) in the presence of substrates.

The $\Delta\Delta G^*_{inact}$ (the change in activation energy of thermal inactivation, i.e. ΔG^*_{wt} - ΔG^*_{mutant}) can be calculated from the first order rate constants for thermal inactivation (k_{inact}) using the Arrhenius equation. The Arrhenius equation states:

$$k = Ae^{-Ea/RT}$$

where k= rate constant, A= frequency factor or pre-exponential factor, Ea= activation energy, R= the gas constant (8.314 J/mol/K) and T=absolute temperature (K).

This equation can be rearranged to give a straight-line equation:

$$\ln k = \frac{-Ea}{R}.\frac{1}{T} + \ln(A)$$

The Arrhenius activation energy (Ea) equates to the activation energy (ΔG^*) as described in transition state theory (illustrated in Figure 3.10), that is, the difference in free energy between substrates and transition state. Hence, the equation can be stated as:

$$\ln k = \frac{-\Delta G^*}{RT} + \ln(A) \qquad \textbf{(equation 1)}$$

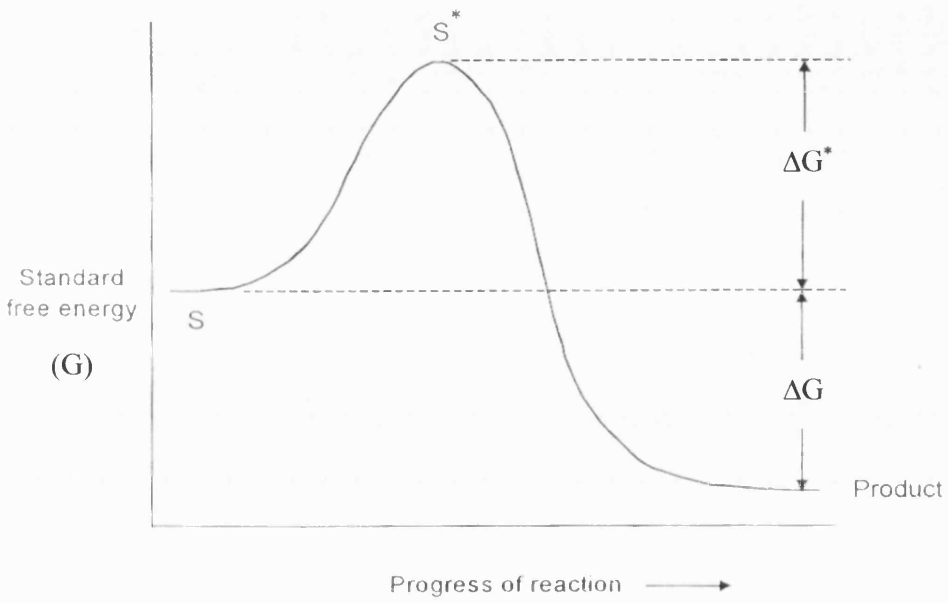


Figure 3.10 **Diagram to illustrate the transition state theory.**
Figure describes the enzymic reaction of S (substrate) to P (product) via a reaction transition state (S^*). The free energy difference between the substrate and the transition state is described as ΔG^*, and the free energy difference between the substrate and the product is described as ΔG.

Hence, the following simultaneous equation can be solved:

$$\ln k_{wt} = \frac{-\Delta G^{*}_{wt}}{RT} + \ln(A) \qquad \textbf{(equation 2)}$$

$$\ln k_{mut} = \frac{-\Delta G^{*}_{mut}}{RT} + \ln(A) \qquad \textbf{(equation 3)}$$

$$\ln k_{wt} - \ln k_{mut} = \left(\frac{-\Delta G^{*}_{wt}}{RT} + \ln(A)\right) - \left(\frac{-\Delta G^{*}_{mut}}{RT} + \ln(A)\right) \qquad \textbf{(eqn2-eqn3)}$$

$$\ln\left(\frac{k_{wt}}{k_{mut}}\right) = \frac{-\Delta G^{*}_{wt}}{RT} + \frac{\Delta G^{*}_{mut}}{RT}$$

$$RT \ln\left(\frac{k_{wt}}{k_{mut}}\right) = \Delta G^{*}_{mut} - \Delta G^{*}_{wt} = \Delta\Delta G^{*}_{inact}$$

$$\Delta\Delta G^{*}_{inact} = RT \ln\left(\frac{k_{wt}}{k_{mut}}\right) \qquad \textbf{(equation 4)}$$

Equation 4 is used to calculate that difference in activation energy between the *Pf*CS wt and the *Pf*CS mutants in both the absence and presence of substrates, and also the change of activation energy caused by the presence of substrate in a single CS.

Table 3.4 shows the $\Delta\Delta G^{*}_{inact}$ values determined from the k_{inact} data in Table 3.3.

	Difference in ΔG^{*}_{inact} of mutants compared to *Pf*CS wild-type		Difference in ΔG^{*}_{inact} between same enzyme (with and without substrate)
Enzyme	$\Delta\Delta G^{*}_{inact}$ in absence of substrates (kJ/mol)	$\Delta\Delta G^{*}_{inact}$ in presence of substrates (kJ/mol)	$\Delta\Delta G^{*}_{inact}$ between absence and presence of substrates (kJ/mol)
*Pf*CS wild-type	-	-	+2.70 (±0.22)* +1.029(±0.14)†
*Pf*CS (-2)	-3.37 (±0.16)*	-2.86 (±0.23)	+3.19 (±0.18)
*Pf*CS (-13)	-1.55 (±0.09)*	-2.01 (±0.13)	+2.21 (±0.10)
*Pf*CS (D113A)	-3.98 (±0.43)†	-4.16 (±0.28)	+1.10 (±0.11)
*Pf*CS (D113S)	-2.35 (±0.24)†	-4.16 (±0.32)	-0.53 (±0.04)

Table 3.4 **Change in activation energies of thermal inactivation for *Pf*CS wt and subunit interface mutants.**
$\Delta\Delta G^{*}_{inact}$ values were calculated from k_{inact} data (Table 3.4) obtained from the thermal inactivations at 100°C (Figures 3.8 and 3.9) using equation 4 derived previously. Two values for *Pf*CS wt are shown, * indicates the values obtained from Figure 3.9(a) incubated at 103°C for the inactivation of the C-terminal mutants. † indicates the values obtained from the thermal inactivation at 100°C (Figure 3.9(b)) for the inactivation of the ionic network mutants.

The data in Table 3.4 show that although the presence of substrates decreases the value of k_{inact} for all *Pf*CSs (except *Pf*CS (D113S)), that is, the enzymes become less susceptible to irreversible thermal denaturation, there is no change to the relative stabilities. The $\Delta\Delta G^{*}_{inact}$ values for the difference between *Pf*CS wild-type and *Pf*CS mutants change by a similar magnitude in both the absence and presence of substrates. The difference between the ΔG^{*}_{inact} values of the same *Pf*CS in the absence or presence of substrate remains relatively comparable for *Pf*CS and all the *Pf*CS mutants. This indicates that the presence of substrates in the thermal inactivation results in the same magnitude of change in

ΔG^{*}_{inact} for each citrate synthase, suggesting that the substrates do not change the effect of the mutation on the citrate synthase enzyme.

3.7 Temperature optima of *Pf*CSs.

The temperature dependence of enzyme activity of *Pf*CS wt and the *Pf*CS mutants, and their temperature optima (T_{opt}), were determined by carrying out 2-minute enzyme assays at different temperatures between 40-90°C. Increased OAA concentrations were used so that this substrate remained at saturating levels despite its thermolability at high temperatures (half-life ~0.8 min at 90°C). Degradation of AcCoA was shown to be insignificant during the time course of the assay. Each measured rate of reaction was corrected for the rate of degradation of DTNB, which increases with elevation in temperature, and which therefore results in an over-estimation of catalytic activity (up to 10% measured activity at 90°C).

Figure 3.11 shows that the T_{opt} for *Pf*CS wild-type is >90°C, which is consistent with previous experimental data [Muir *et al.*, 1995]. All of the *Pf*CS mutants display a decrease in T_{opt} compared to *Pf*CS wt indicating that the disruption in interactions at the subunit interface have affected the thermoactivity of the *Pf*CS wt enzyme. As shown in Table 3.2, the V_{max} values of the *Pf*CS wt and the *Pf*CS mutants at 55°C (pH 8) are comparable, which is also demonstrated in Figure 3.11 where the specific activities at 55°C (pH 7) are very similar to each other. It should be noted that the temperature-dependence of catalytic activity (Figure 3.11) was determined at pH 7; that is, under the same conditions as those in which the thermal inactivations were carried out (Figures 3.8 and 3.9). For this reason, the

specific activities given in these figures are approximately half of those given in Table 3.2, which were determined at pH 8, nearer the pH optimum for citrate synthase.

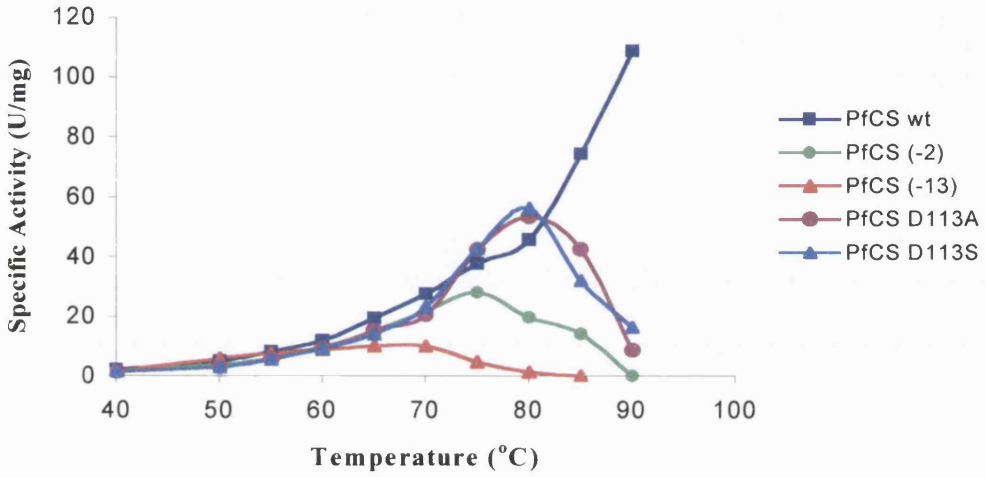


Figure 3.11 **Temperature dependence of the catalytic activities of *Pf*CS wt and *Pf*CS mutants.**
Enzyme assays were carried out at pH 7.0 over a range of temperatures, 40-90°C, and initial rates were measured over 2 minutes in each case. Substrate concentrations of 0.14mM acetyl-CoA and 1.25mM OAA were used in all cases, except *Pf*CS (-13) where 1.2mM acetyl-CoA and 2.5mM OAA were used.

Both ionic network mutants, *Pf*CS (D113A) and *Pf*CS (D113S), have a T_{opt} of ~80°C, which is significantly lower than the T_{opt} of *Pf*CS wild-type, >90°C. Figure 3.11 also shows that the ionic network mutants have similar specific activities to the *Pf*CS wild-type up until their T_{opt} at 80°C. This suggest that the mutants have no significant change in their catalytic activity compared to *Pf*CS wt, which correlates to the kinetic data, but undergo a temperature induced conformational change which results in loss of activity. The C-terminal deletion mutants also demonstrate this, although these mutants show a loss of activity at lower temperatures than the *Pf*CS wt and the ionic network mutants. *Pf*CS (-13) is seen to be the most temperature sensitive with a T_{opt} of ~65°C, whereas *Pf*CS (-2) T_{opt} of ~75°C.

The implication that the mutants have no great change in catalytic activity but are sensitive to a temperature induced conformational change that will affect the activity at elevated temperatures can be illustrated in an Arrhenius equation plot. The data in Figure 3.11 can be plotted as a straight-line Arrhenius plot using equation 1 as in Section 3.6:

$$\ln k = \frac{-\Delta G^*}{RT} + \ln(A)$$

If it is assumed that A and ΔG^*_{cat} (the activation energy of the citrate synthase catalytic reaction) are approximately constant values over a moderate range of temperature (~50K), then $\ln k$ can be plotted against $1/T$ to give a straight line of slope equal to $-\Delta G^*_{cat}/R$ and

the intercept on the ordinate gives $\ln A$. Figure 3.12 shows the temperature optima data (Figure 3.11) plotted using the linear form of the Arrhenius equation.

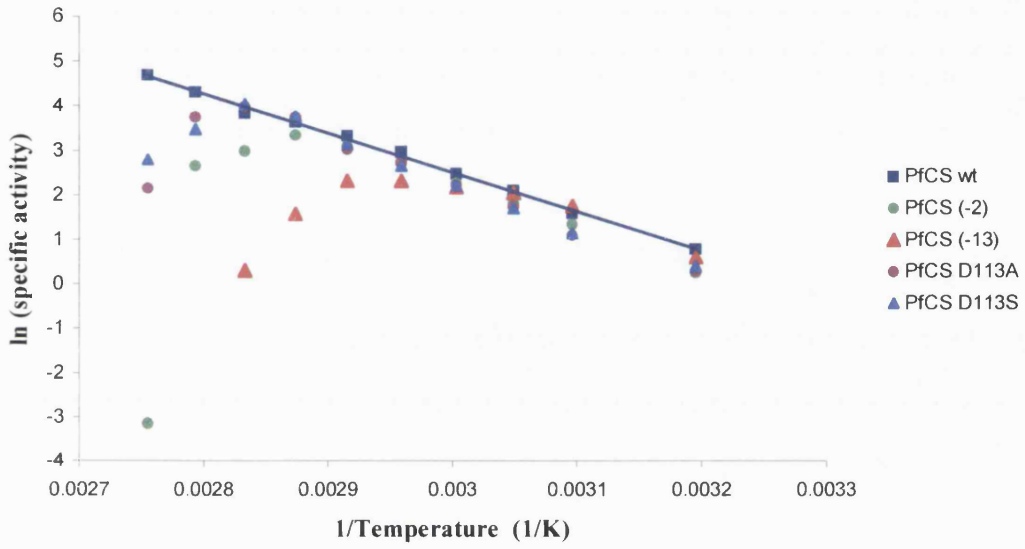


Figure 3.12 **Arrhenius plot of *Pf*CS wt and *Pf*CS subunit interface mutants.** Temperature dependence data plotted as Arrhenius plot. Slope of line is equal to $-\Delta G^{*}_{cat}/R$. Line is fitted to *Pf*CS wt data for illustrative purposes.

The Arrhenius plot (Figure 3.12) shows that there is no significant change in the activity of the *Pf*CS mutants compared to the *Pf*CS wt at lower temperatures. The differences in the activation energies of the catalytic reactions (ΔG^{*}_{cat}) become increasingly significant with elevated temperatures. Table 3.5 summarises the ΔG^{*}_{cat} values calculated from the linear portion of the Arrhenius plot (Figure 3.12). These data strongly argue that there is no

significant difference in the catalytic power of the *Pf*CS subunit interface mutants compared to the *Pf*CS wt until increased temperatures prevent catalytic activity. This suggests that the mutations at the dimer interface region make the *Pf*CS susceptible to temperature-induced conformations, which result in the loss of activity at increased temperatures.

Enzyme	ΔG^{*}_{cat} (kJmol)
*Pf*CS wild-type	73. (±1.5)
*Pf*CS (-2)	85 (±1.5)
PfCS (-13)	71 (±12)
PfCS (D113A)	88 (±3)
*Pf*CS (D113S)	87 (±3)

Table 3.5 **Activation energies of the catalytic reactions of *Pf*CS wt and mutants.** Activation energies are calculated from the linear portions of the Arrhenius plot (Figure 3.12).

3.8 Discussion

The *Pf*CS wt and each of the 4 mutants were successfully expressed in the high expression vector pREC7/*NdeI*. Purification of the citrate synthases was aided by their expression in a citrate synthase negative strain of *E.coli*, MOB154. The absence of any mesophilic citrate synthase and the thermostability of the expressed citrate synthases enabled the enzymes to be partially purified by a simple heat treatment. This heat step (85°C for the *Pf*CS wild-type and 80°C for the mutants) removed a high proportion of the protein from the cell extract (Figure 3.5) and, even though this heat step was above the T_{opt} for the C-terminal deletion mutants (Figure 3.11), there was no significant loss of citrate synthase activity (Table 3.1).

The Dye Matrex Red A column achieved the final purification in each case (Figure 3.5 and Table 3.1). Dye Matrex Red A has been used previously for the purification of eukaryotic and bacterial acyl-CoA utilizing enzymes, including citrate synthase, on the basis that the dye has a preferential affinity for NADP-linked enzymes and that Coenzyme-A shares common structural features with NADP [James *et al.*, 1994]

The intention of the mutations at the subunit interface of *Pf*CS is to study the role of inter-subunit interactions specific to the thermostability of this enzyme. As this study is concerned with determining the molecular basis of thermostability, it is important that only regions of the protein concerned with thermostability are altered. This is particularly important in the citrate synthase enzyme as the active sites lie between the two subunits,

requiring residues from both subunits for activity. Any large disruption between the subunits could subsequently affect the activity of the enzyme. To determine whether a mutated enzyme has significantly altered active-site structure compared to the wild-type enzyme it is useful to compare the kinetic parameters of the enzymes studied.

The kinetic data displayed in Table 3.2 show that the K_m, V_{max} and k_{cat} values are consistent between each *Pf*CS, apart from the values obtained for the *Pf*CS (-13) C-terminal deletion mutant. This suggests that the ionic network mutations, D113A and D113S, and the –2 deletion mutation do not greatly affect the active site of *Pf*CS at 55°C as the K_m and V_{max} values are comparable at this temperature. The *Pf*CS (-13) displays K_m values that are greatly increased compared to the *Pf*CS wt (~10-fold increase), although the V_{max} values suggest that there is no change in the overall catalytic activity. The increase in the K_m values for this enzyme means that the concentrations of both OAA and AcCoA need to be increased so that the enzyme may be completely saturated. If the enzyme is not saturated then the apparent activity will be lower than the V_{max} value. This is illustrated in Table 3.1 where the purified fractions were assayed before the kinetic parameters were determined so the substrate concentrations were non-saturating and the specific activity appears lower than the V_{max} values later obtained for this enzyme.

These changes in the kinetic parameters of *Pf*CS (-13) suggest that the activity of this citrate synthase has been affected by the complete deletion of the C-terminal arm region. Structural data suggest that this region of the enzymes may be involved in the association of the subunits [Russell *et al.*, 1994], so if this region is removed and the subunits no longer

associate as intimately, then the activity of the enzyme would be compromised. This does not appear to apply to the *Pf*CS (-2) where only the two terminal arginine residues are deleted as illustrated in Figure 3.2. These arginine residues are basic in nature and the penultimate arginine (375) binds ionically to a glutamate residue (48) on the other subunit. The removal of these residues in *Pf*CS (-2) does not appear to affect the kinetic properties of this enzyme with respect to the *Pf*CS wild-type implying that this mutation does not affect the active site structure of *Pf*CS.

Thermal inactivation data show that both the C-terminal deletion mutants and the ionic network mutants are considerably less thermostable than *Pf*CS wild-type when incubated in the absence of substrates (Figure 3.8, data produced by Arnott (1999)). For the C-terminal deletion mutants it seems that removing the two terminal arginines causes a greater reduction in thermostability of *Pf*CS than removing the entire 13-residue arm region even though the -13 mutation is more detrimental to the catalytic activity. These data may suggest that the loss of the ionic interactions, when the terminal arginine residues are removed, results in a gain of conformational entropy (more flexibility in this region now that the arm is not 'anchored' to the other subunit). This flexible region may now be more susceptible to thermal denaturation. The complete removal of this C-terminal flexible region (the –13 mutation) may stabilize the enzyme, even though this causes a change in the catalytic properties at 55°C. There is evidence that indicates the reduction of flexibility in a protein may contribute to an increase in thermostability [Daniel *et al.*, 1996].

In the case of the ionic network mutants, the thermal inactivation data show that both *Pf*CS (D113A) and *Pf*CS (D113S) have an increased rate of thermal inactivation compared to the *Pf*CS wild-type with *Pf*CS (D113A) being the most susceptible to irreversible denaturation. This may reflect the ability of the polar serine side chain to participate in a hydrogen bonding interaction with the remaining charged residues of the network which the smaller, hydrophobic alanine residue is unable to do, according to the initial hypothesis (see previous work, Section 3.2). Pappenberger *et al.* (1997) showed that the disruption of the ionic network in the D-glyceraldehyde-3-phosphate dehydrogenase from *Thermotoga maritima* resulted in a decrease in free energy of activation for thermal unfolding of about 4kJ/mol at 100°C. This is comparable with results shown in Table 3.4 where the *Pf*CS ionic network mutants show a decrease in the activation energies of thermal inactivation of 2.3-4.2 kJ/mol about 100°C in both the absence and presence of substrates. Thermal inactivation in the presence of substrates (Figure 3.9) however, shows that there is a lower rate of irreversible denaturation for all of the *Pf*CS mutants, except *Pf*CS (D113S) suggesting that the presence of substrate provides a degree of stability at high temperatures. The effect of the presence of substrates is more apparent in the case of the ionic network mutants (Figure 3.9(a)), where there is no significant difference between *Pf*CS (D113A) and *Pf*CS (D113S). These effects are not surprising in the case of citrate synthase as the binding of substrates requires residues from both subunits, so the substrates may provide an extra interaction between the two monomers and aid in maintaining the integrity of the dimer.

Although the effects of the presence of substrates are visible from the thermal inactivation data (Figures 3.8 and 3.9), Table 3.4 shows that the overall relative stabilities remain unaltered compared to the *Pf*CS wild-type. These data also show that the presence of substrates causes the same degree of stabilisation in each citrate synthase. Hence the substrates have little effect on the change in the thermostability of the *Pf*CS induced by the mutations, but have a general stabilising effect on all of the *Pf*CSs.

The temperature dependence data obtained, illustrated in Figure 3.11, shows a reduction in T_{opt} compared to *Pf*CS wild-type for all PfCS mutants; 70°C for *Pf*CS (-2), 65°C for *Pf*CS (-13) and 80°C for both ionic network mutants compared to >90°C for *Pf*CS wt. This figure also shows that at 90°C there is no *Pf*CS (-2) or *Pf*CS (-13) activity remaining which contradicts the thermal inactivation data that shows both enzymes can survive a five minute incubation at 100°C without complete irreversible denaturation. This implies that the loss of activity after the T_{opt} (Figure 3.11) is not solely due to irreversible denaturation [Daniels *et al.*, 2000]. The same applies for the ionic network mutants as they have very little activity remaining at 90°C (Figure 3.11), but have residual activity after an incubation at 100°C for >5 minutes (Figures 3.8 and 3.9). These data strongly imply that the loss of activity after the T_{opt} for each *Pf*CS can be due to a conformational change that results in the loss of activity, rather than complete irreversible denaturation of the protein. This conformational change may be associated with a dissociation of the dimers since this is the region of the enzyme that has been mutated. As the active site of this enzyme requires residues from both monomers for activity, then a dissociation of the dimers would result in a loss of activity. This would explain the difference of the T_{opt} values between the *Pf*CS

mutants as each mutation could have caused a different degree of instability within the dimer interface region of the protein. This theory is consistent with the data derived from the Arrhenius plot (Figure 3.12 and Table 3.5), which show no significant difference in the activation energy of catalysis for each *Pf*CS at lower temperatures. Change in activation energies only occurs at higher temperatures when a thermally-induced change in the *Pf*CS mutants causes a loss of activity.

In conclusion, characterisation of the *Pf*CS subunit interface mutants has shown that it is possible to change the thermostability of the *Pf*CS wild-type without, in all but one case, significantly changing the catalytic abilities of the enzyme. There is strong evidence provided here to demonstrate the role of both the ionic network and the C-terminal arm region in the hyperthermostability of the *Pf*CS. As reviewed in the Introduction, ionic interactions, especially ionic networks, are now thought to be a common feature of enzymes that function near to or above 100°C.

CHAPTER 4

Construction of inter-subunit ionic network mutants of *Thermoplasma acidophilum* citrate synthase.

4.1 Introduction

Thermoplasma acidophilum is a thermophilic Archaeon that grows optimally at 55-60°C, pH 1-3. The gene encoding citrate synthase from *T.acidophilum* has been cloned and over-expressed in *E.coli*, and the recombinant enzyme purified [Sutherland *et al.*, 1990; 1991]. The open (ligand-free) crystal structure has been determined to a resolution of 2.5Å [Russell *et al.*, 1994]. The *Thermoplasma acidophilum* citrate synthase (*Ta*CS wt) shows optimal activity at ~75°C, compared to the recombinant *Pf*CS wt which has a temperature optimum >90°C [Arnott *et al.*, 2000].

The *Ta*CS wt amino acid sequence is 42% identical (62% similar) to the citrate synthase amino acid sequence from *Pyrococcus furiosus*, and there is a very high degree of structural homology. As for the *Pf*CS wt, the *Ta*CS wt consists of two monomers (each of 43kDa) [Sutherland *et al.*, 1991], composed of 16 α-helices distributed between a large and a small domain, and the interface of the two subunits is composed of an 8 α-helical sandwich of 4 antiparallel pairs of helices (F,G,M and L).

As previously discussed, structural trends with increasing temperature were identified when comparing the crystal structures of citrate synthase spanning the biological temperature range (10-100°C) and, as described in Chapter 3, this investigation is primarily concerned with structural changes at the dimer interface of the *Ta*CS enzyme. The *Ta*CS wt has an increase in hydrophobic interactions at the subunit interface compared to the citrate synthase from the mesophilic pig (37°C); however, there is in addition one ionic interaction in this region compared to two ionic interactions found at the interface of the CS from *Sulfolobus solfataricus* (85°C) and the five-membered ionic network found in this region in the CS from *P.furiosus* (100°C) (see Figure 3.2).

The previous chapter (Chapter 3) discusses the effects of disrupting inter-subunit contacts to investigate their role in the structural basis of thermostability in the *Pf*CS. As previously discussed (see Section 3.1), the additional ion pairs found at the subunit interface of *Pf*CS wt are located in Helix G and Helix M that make up the four central helices of the eight helical sandwich. Figures 4.1(a) and 4.1(b) illustrate the positioning of the ionic network of *Pf*CS wt compared to the equivalent region in the *Ta*CS wt.

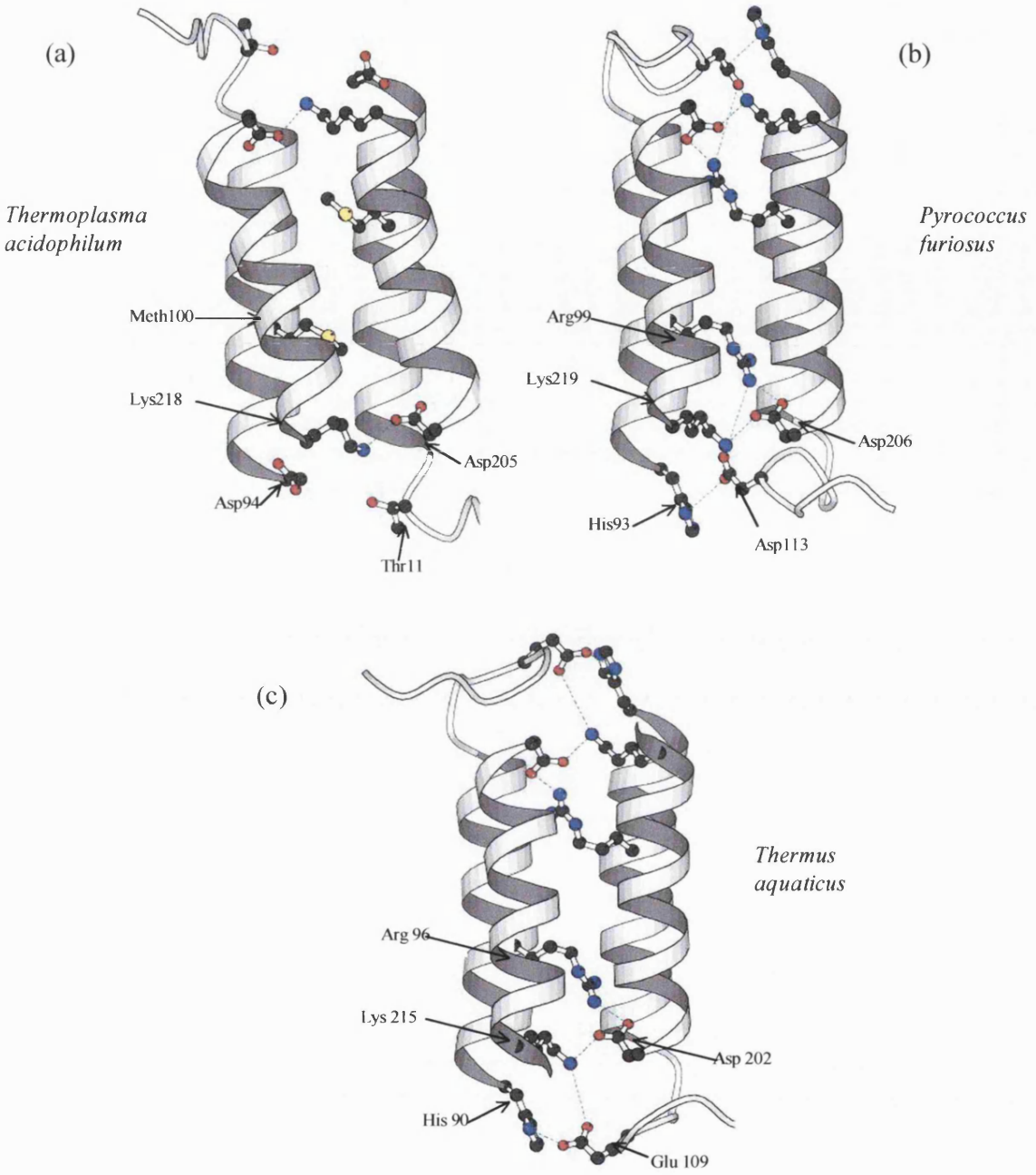


Figure 4.1 **Comparison of intersubunit central helices (G and M) of citrate synthases from *T.acidophilum*, *P.furiosus* and *T.aquaticus*.** Ribbon diagrams show helices G and M of both subunits at the subunit interface. Equivalent residues that make up the ionic network in *Pf*CS are labeled in each case. This figure shows that while *Ta*CS has only one ionic bond, the thermophilic *Taq*CS has four ionic bonds.

The aim of the investigation in this chapter is to introduce a number of ionic interactions at the subunit interface of *Ta*CS wt by site-directed mutagenesis with the final aim of constructing a five-membered ionic network equivalent to that found at the dimer interface of *Pf*CS wt. Only a few examples have been reported where such a modification has been successfully achieved. Nemeth *et al.* (2000) created an inter-subunit ion pair that could join existing ion clusters at the dimer interface of 3-isopropylmalate dehydrogenase from *E.coli* to mirror the existing ionic interactions in the thermophilic *Thermus thermophilus* enzyme. They found that this substitution resulted in an increase in melting temperature of 2-4°C of the mesophilic enzyme. Lebbink *et al.* (1999) investigated the role of an 18-residue ionic network that is present in the glutamate dehydrogenase (GDH) of the hyperthermostable Archaeon *P.furiosus* by introducing four new charged residues into the slightly less hyperthermostable glutamate dehydrogenase of the bacterium *Thermotoga maritima,* thereby to mirror the ionic network in the *P.furiosus* enzyme. Amino acid substitutions were introduced as single mutations as well as in several combinations. This study showed that some single amino acid substitutions resulted in a decrease in resistance towards thermal inactivation, but a combination of destabilizing single mutations restored stability in most cases. In one case a combination of three destabilizing mutations increased the thermostability of the *T.maritima* GDH i.e. increased the half-life of inactivation at 85°C by 30 minutes, increased the temperature optimum by 3°C and increased the melting temperature by 0.5°C. This is good evidence of the need for balanced charges at subunit interfaces and high cooperativity between the members of the ionic network.

Figure 4.2 summarises the position and substitution of the mutations required to introduce an ionic network, equivalent to the five-membered ionic network present at the subunit interface of *Pf*CS wt, into the *Ta*CS wt. In this figure each mutation is named RAM3, RAM4, RAM5 or RAM6, which correspond to the primer used to introduce these mutations. The primers are discussed further in Section 4.2.2. *Ta*CS wt has one ionic bond present at the subunit interface, Lys218-Asp205, which corresponds to one of the ionic bonds in the ionic network of *Pf*CS wt, Lys219-Asp206, also shown in Figure 4.1.

4.1.1 Methionine100-Aspartate, M100R (RAM3).

The first substitution in *Ta*CS wt is Meth100 to an arginine residue, M100R (RAM3). This should form an ionic interaction with Asp205, although this interaction is predicted to form a bond of ~3.6Å compared to the equivalent bond in *Pf*CS wt (2.9Å). One problem with the M100R substitution is its close proximity to Lys218 (2.1Å) compared to a distance of 4.9Å between Lys219 and Arg99 in *Pf*CS wt. This may cause a clash between the positively charged groups in the *Ta*CS, which may result in an unfavourable structural change at the subunit interface.

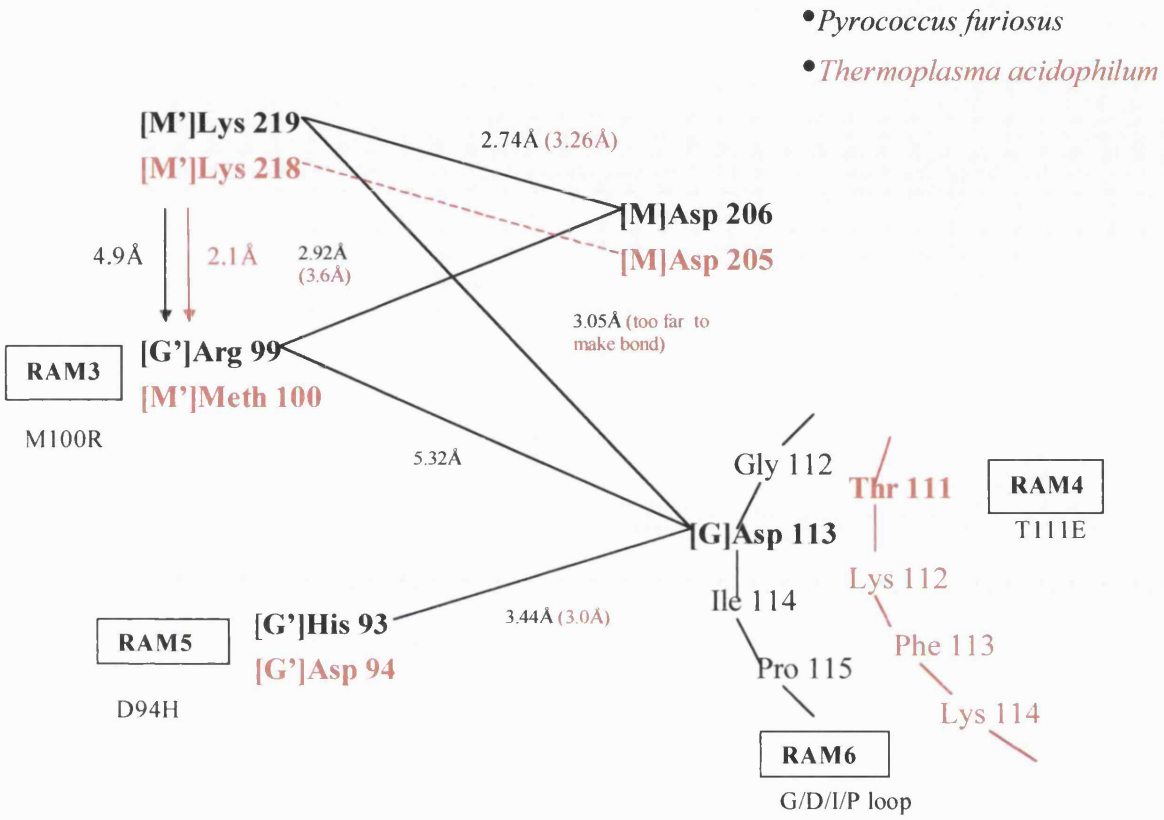


Figure 4.2 **Schematic comparison of ionic network regions of *Ta*CS wt and *Pf*CS wt.** *T.acidophilum* residues are shown in red and *P.furiosus* residues are shown in black. The lengths of the ionic bonds in the *P.furiosus* ionic network are shown in black, and in red for *T.acidophilum* where possible. Each change of residue between the two citrate synthases is named corresponding to the primer used to make this mutation in the *Ta*CS wt (RAM3, RAM4, RAM5 and RAM6).

4.1.2 Threonine111-Glutamate, T111E (RAM4)

The second substitution made to *Ta*CS wt is the substitution by Thr111 with a glutamate residue, T111E (RAM4), (Figure 4.2). Figure 4.1(b) shows that the Asp113 residue of *Pf*CS wt is positioned on a four-residue loop that is not present in the *Ta*CS wt. The nearest equivalent residue in the *Ta*CS wt structure is Thr111 (labeled on Figure 4.1(a)). If the Thr111 in *Ta*CS wt were to be directly substituted with an aspartate residue (when building the *Pf*CS wt ionic network into *Ta*CS wt) it would be impossible for this residue to form any ionic bonds with other members of the ionic network because of the distance of this residue from the rest of the ionic network.

The ionic network present in the citrate synthase from the thermophilic bacterium, *Thermus aquaticus* (*Taq*CS wt) suggests a solution to this problem. Figure 4.1 (c) shows a model of the central four helices of the subunit interface of *Taq*CS wt. The crystal structure of this enzyme has not yet been determined so a comparison can only be made to the sequence-based model of this protein [Nordberg-Karlsson, E., Crennell, S.J., Danson, M.D. & Hough, D.W., unpublished data]. As the *Taq*CS wt has the greatest sequence identity with *Pf*CS wt (47%) then the latter crystal structure has been used as a basis for the modeling.

The sequence-based model of the *Taq*CS wt shows that this enzyme is structurally homologous to the *Ta*CS wt, *Sulfolobus solfataricus* citrate synthase (*Ss*CS wt) and the *Pf*CS wt. *Thermus aquaticus* has a growth optimum of ~70°C, so it lies between the growth optima for *T.acidophilum* (55°C) and *P.furiosus* (100°C). The *Taq*CS wt has a

temperature optimum of 80°C. As the increased thermostability of citrate synthase appears to be correlated with an increase in ionic interactions, particularly at the subunit interface, it would be expected that the *Taq*CS wt subunit interface would contain between one ion pair (as in *Ta*CS) and a five-membered ionic network (as in *Pf*CS). Figure 4.1 suggests that this hypothesis may be true in this case. The *Taq*CS wt model indicates that there may be five residues forming four ionic bonds between the four central helices of the subunit interface, although this in an increase in ionic bonds compared to the *Ss*CS wt, that has a temperature optimum of ~90°C, which has two ionic bonds. The simple ionic network of *Taq*CS wt is equivalent to the ionic network found in the *Pf*CS wt with the exception of the Asp113 in *Pf*CS wt which is Glu109 in *Taq*CS wt. Interestingly, this is the aspartate residue that is substituted in an attempt to decrease the thermostability of *Pf*CS wt in Chapter 3. This is interesting as it suggests that the increase of ion pairs at the subunit interface with thermostability is not a feature unique only to the archaeal citrate synthases.

Therefore, a glutamate residue is used to substitute the Thr111 in the *Ta*CS wt in an attempt to mirror the Asp113 residue in the *Pf*CS wt. The glutamate residue may be used in this case as this residue is one methylene group longer than the aspartate residue, hence, this may make up the difference in distance from the other members of the ionic network enough to allow the formation ionic interactions.

Figure 4.2 illustrates that the substitution of Thr111 to a glutamate may cause a clash with the negatively charged Asp94. Also, this substitution may still be insufficient to make up

the distance so an alternate mutation was also made, RAM6, which is the insertion of the loop region that is present in *Pf*CS wt (Section 4.1.4).

4.1.3 Aspartate94-Histidine, D94H (RAM5)

The third mutation is to substitute Asp94 with a histidine residue, D94H (RAM5), in an attempt to create an ionic bond with the residue that mirrors the Asp113 in the *Pf*CS wt. This mutation is relatively exposed to the solvent so, theoretically, a mutation at this site should not cause extensive disruption to the overall protein structure.

4.1.4 Four-residue loop insertion, G/D/I/P (RAM6)

The final mutation is an alternative to the T111E (RAM4) mutation, discussed in Section 4.1.2. Figure 4.1 shows that the Asp113 residue in the *Pf*CS wt is present on a four-residue loop, which is absent from the *Ta*CS wt. This loop appears to be responsible for positioning the Asp113 residue near to the subunit interface, enabling the formation of ionic interactions. Although the T111E (RAM4) mutation attempts to position a glutamate residue in the equivalent position, it may prove more successful to insert the complete four-residue loop into the *Ta*CS wt.

The loop in *Pf*CS wt is made up of Gly112, Asp113, Ile114 and Pro115 (G/D/I/P). These four residues are substituted into the *Ta*CS wt in place of the Phe113 and Lys114 residues.

Preliminary study of the *Ta*CS wt structure has suggested that there is sufficient room in this region for the loop to be inserted.

The four mutations designed were introduced into the *Ta*CS wt structure in different combinations to try to minimize any disruption caused by introducing residues with clashing charges. A list of all the *Ta*CS wt ionic network mutants are summarized in Section 4.2.2.

Preliminary modeling of the mutations, discussed in Chapter 5 (Section 5.6), into the *Ta*CS wt revealed no major problems; hence, this chapter describes the methods used for the construction of all the *Ta*CS ionic network mutants and the purification of these enzymes to a degree sufficient for kinetic and thermal characterization. The characterization of the *Ta*CS wt and the mutants is described and discussed in Chapter 5.

4.2 Methods

4.2.1 Nested PCR

As it was necessary to use a PCR method for site-directed mutagenesis that was able to introduce a mutation into a gene and amplify it in sufficient quantity for cloning in relatively few steps, a two-step nested-PCR (or megaprimer PCR) method was selected. The principle for this nested-PCR method is shown in Figure 4.3.

The 1st round PCR uses a sense primer (RAM1) containing an *NdeI* restriction site, and a mutagenic primer (RAM3, RAM4, RAM5 or RAM6) to generate a fragment with the desired mutation. In the case of mutating and cloning wild-type and mutant citrate synthases of *T.acidophilum*, this fragment is expected to be ~300 bases in length. The product of this PCR is used in the 2nd round PCR along with an antisense primer (RAM2) containing a *KpnI* restriction site to generate the full length gene (~1.3kb) for purification and cloning into the pREC7/*NdeI* expression vector. All primers used are discussed in detail in Section 4.2.2. The exact protocol for this nested-PCR method is described, along with the results, in Section 4.4.

This method of site-directed mutagenesis is very useful as it enables the construction of *Ta*CS genes containing multiple combinations of the mutations by using different mutagenic primers and using previously constructed *Ta*CS mutants as DNA templates to introduce further mutations.

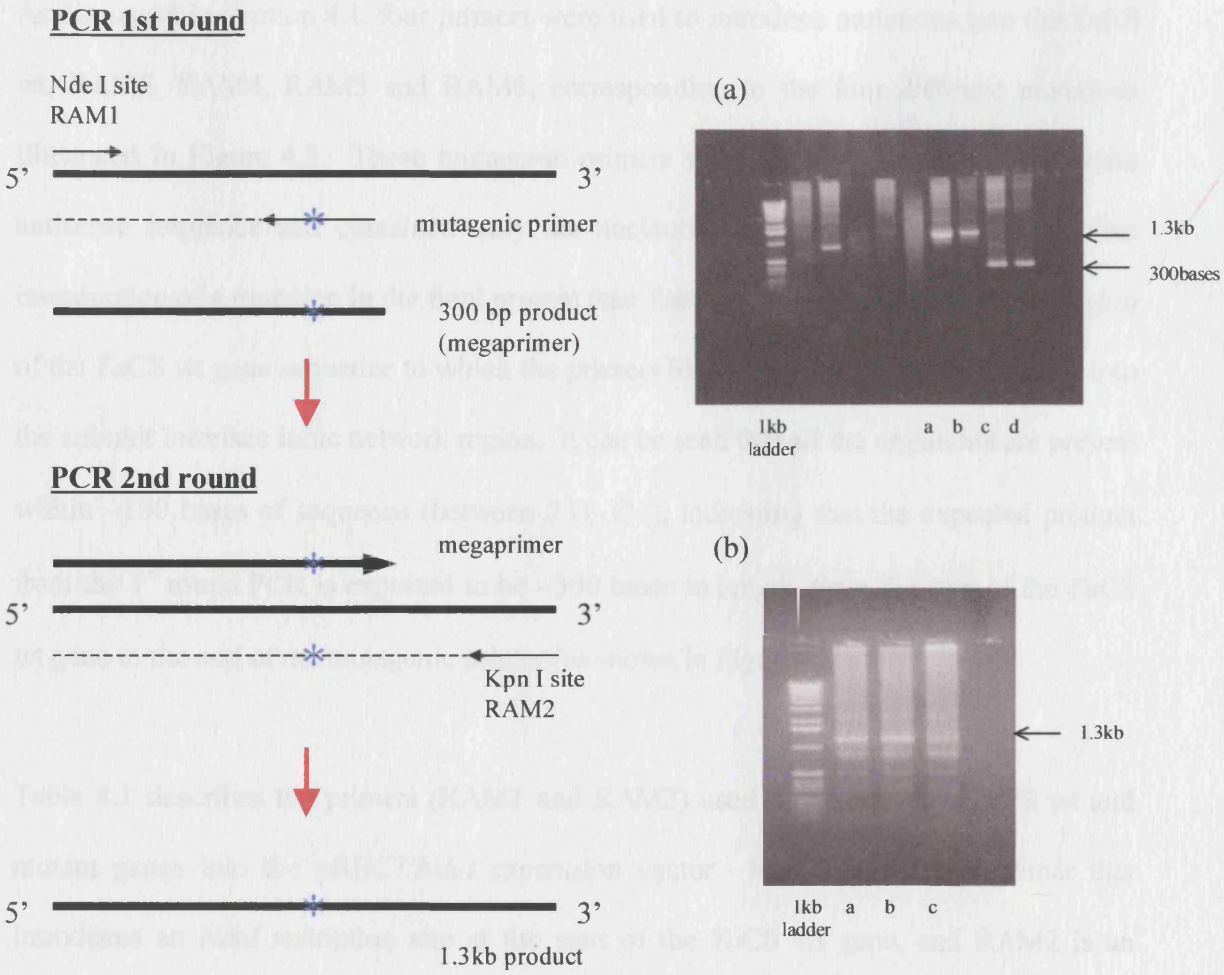


Figure 4.3 **Nested PCR method for the construction of *Ta*CS ionic network mutants.** Diagram shows the two stages of the nested PCR protocol. The 1st round PCR uses a sense primer (RAM1) containing an *NdeI* restriction site and a mutagenic primer (RAM3, RAM4, RAM5 or RAM6). The product of this PCR becomes the megaprimer in the 2nd round PCR along with an antisense primer (RAM2) containing a *KpnI* restriction site. The gels illustrate the expected products of this PCR method. Gel (a) shows that in the 1st round PCR a product of ~300 bases is formed (lanes c and d). Lanes a and b on gel (a) show the 1.3kb product (entire *Ta*CS wt gene) if RAM1 and RAM2 are used for cloning purposes. Gel (b) shows the 1.3kb product of the 2nd round PCR which is expected to be the full length *Ta*CS gene containing the introduced mutation. The DNA templates and products are shown as a thick line. The mutation introduced is shown as ✳.

4.2.2 Primers

As discussed in Section 4.1, four primers were used to introduce mutations into the TaCS wt, RAM3, RAM4, RAM5 and RAM6, corresponding to the four different mutations illustrated in Figure 4.2. These mutagenic primers were designed to correspond to the antisense sequence and contained only the nucleotide replacements necessary for the introduction of a mutation in the final protein (see Table 4.1). Figure 4.4 shows the region of the *Ta*CS wt gene sequence to which the primers bind when introducing mutations into the subunit interface ionic network region. It can be seen that all the mutations are present within ~100 bases of sequence (between 231-321), indicating that the expected product from the 1st round PCR is expected to be ~300 bases in length, from the start of the *Ta*CS wt gene to the end of the mutagenic primer (as shown in Figure 4.3).

Table 4.1 describes the primers (RAM1 and RAM2) used for cloning the *Ta*CS wt and mutant genes into the pREC7/*NdeI* expression vector. RAM1 is a sense primer that introduces an *NdeI* restriction site at the start of the *Ta*CS wt gene, and RAM2 is an antisense primer that introduces a *KpnI* restriction site at the end of the gene. The positioning of the restriction site within the primer is illustrated in Table 4.2; sequence to the left of this restriction site (5') is necessary for the action of the restriction enzyme. Sequence to the right of the restriction site is identical to the *Ta*CS wt gene for binding. As there are no mutations introduced into these regions of the *Ta*CS wt gene, then RAM1 and RAM2 can also be used for cloning the *Ta*CS ionic network mutants. As shown in Figure 4.3, RAM1 is the sense primer used in the 1st round PCR along with one of the mutagenic

primers (RAM3, 4, 5 or 6). RAM2 is the antisense primer used in the 2nd round PCR with the megaprimer product of the 1st round PCR.

Primer name	**Primer sequence**	**Mutation**
RAM 1	*NdeI* site 5'- GGA ATT CCA TAT GCC AGA AAC TGA AGA AAT TAG -3'	*NdeI* site at start of gene
RAM 2	*KpnI* site 5'- CGG GGG GTA CCT CAC TTT CTT TCA GCG ATG G -3'	*KpnI* site at end of gene
RAM 3	5'- GC AAC CGC GGC CCT CTG CAT TGC CAC AGC -3'	M100R
RAM 4	5'- CCA CTT GAA CTT TTC CTC CGA TGC TGC C -3'	T111E
RAM 5	5'- G CAT TGC CAC AGC ATG CGA TTC CCT TGG C -3'	D94H
RAM 6	5'- CGT ATC CTT GTT CCA CGG AAT GTC GCC CTT TGT CTC CGA TGC -3'	G/D/I/P loop

Table 4.1 **Primer sequences.**
RAM1 and RAM2 primers contain restriction sites which are underlined and in red. The sequence at the 5' end relative to the restriction sites is sequence preferred by the restriction enzyme in both cases. RAM3, RAM4, RAM5 and RAM6 are the mutagenic primers and are shown in colours corresponding to the colours shown in Figure 4.4. In the mutagenic primers the base change is shown in red and the corresponding codon changed is underlined. These primers are named as in Figure 4.2

Figure 4.4 **Binding positions of the mutagenic primers on the *Ta*CS wt gene.** All the mutagenic primers (RAM3, RAM4, RAM5, and RAM6) bind within a region of ~100 bases on the *Ta*CS wt gene. The colours of the primers correspond to the colours shown in Table 4.1. In each case the change in sequence of the primer is underlined and the change in amino acid is shown in the corresponding colour on the protein sequence.

Table 4.2 describes the ten *Ta*CS ionic network mutants created using different combinations of the mutations shown in Figure 4.3 and Table 4.1. For ease of discussion, the names given to the *Ta*CS mutants directly correspond to the number of the primer used for the mutation that the *Ta*CS mutant contains i.e. RAM3 is used to insert the M100R mutation so the mutant containing this mutation is called *Ta*CS 3sm (the **sm** indicating that it is a single **m**utation).

*Ta*CS mutant name	Mutations
*Ta*CS 3sm	M100R
*Ta*CS 4sm	T111E
*Ta*CS 5sm	D94H
*Ta*CS 3/4dm	M100R/T111E
*Ta*CS 4/5dm	T111E/D94H
*Ta*CS 3/4/5tm	M100R/T111E/D94H
*Ta*CS 6sm	G/D/I/P
*Ta*CS 6/3dm	M100R/G/D/I/P
*Ta*CS 6/5dm	D94H/G/D/I/P
*Ta*CS 3/5/6tm	M100R/D94H/G/D/I/P

Table 4.2 ***Ta*CS ionic network mutant nomenclature.**
The name given to each *Ta*CS mutant is shown in the left hand column. Each number in the mutant name corresponds to the primer used for each mutation within the mutant (shown in right hand column). The sm, dm and tm correspond the the number of mutations in each mutant i.e. single mutation, double mutation and triple mutation.

4.3 Cloning of *Ta*CS wt from pUC19 into pREC7/*NdeI*

The *Ta*CS wt gene was previously cloned and expressed in the pUC19 expression vector [Sutherland *et al.*, 1990, Sutherland *et al.*, 1991]. As the *Pf*CS wt and previous *Pf*CS mutants were expressed in the higher expression vector pREC7/*NdeI* (see Chapter 3), it was decided to clone the *Ta*CS into this vector for an increased level of expression and to keep all the citrate synthase expressions consistent for comparative purposes.

The *Ta*CS wild-type gene was amplified from the p*Ta*CS19 construct using the standard PCR protocol (Chapter 2) with the sense primer (RAM1) containing the *NdeI* restriction site and the antisense primer (RAM2) containing the *KpnI* restriction site (Table 4.1). Figure 4.5 shows a 1.3kb PCR product that is consistent with the expected size of the *Ta*CS wt gene [Sutherland *et al.*, 1990].

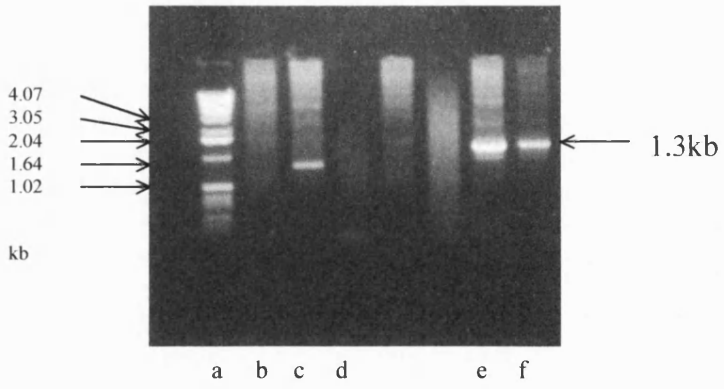


Figure 4.5 **PCR amplification of the *Ta*CS wt gene**
Lane a contains DNA marker (1kb). Lanes b, c and d show the controls for the PCR experiment. Lane b is single primer control for RAM1, and lane c is the single primer control for RAM2. Lane d is the double primer control in the absence of template DNA. Lanes e and f show the PCR product with varying concentrations of template DNA (the reaction for lane f contains half the concentration of template DNA compared to reaction for lane e).

The 1.3kb PCR product was purified from the 1% agarose gel using the QIAEX II kit and digested with *NdeI* and *KpnI* restriction enzymes for 3h. A preparation of the pREC7/*NdeI* expression vector was also digested for 3h with *NdeI* and *KpnI*. Figure 4.6 shows the digested PCR product and the digested pREC7/*NdeI* vector on a 1% agarose gel.

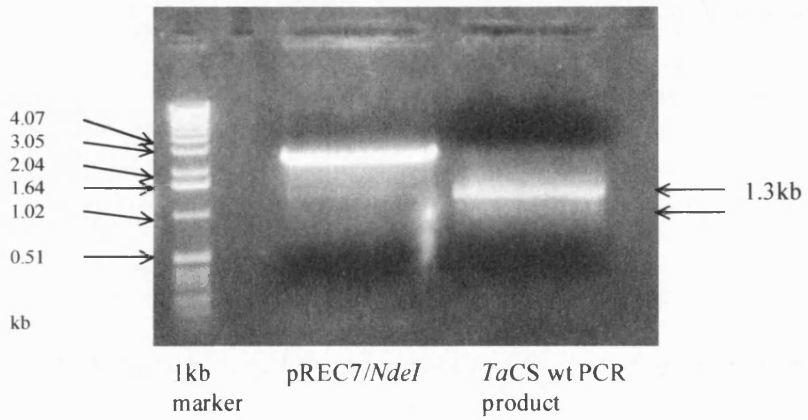


Figure 4.6 **Double digests of pREC7/*NdeI* and *Ta*CS wt PCR product for ligation.**
The gel shows the pREC7/*NdeI* expression vector and the *Ta*CS wt PCR product double digested with *NdeI* and *KpnI* for ligation.

The purified PCR product and vector were ligated using the Rapid DNA ligase kit (using T4 DNA ligase) according to the recommended protocol (see Materials and Method). The successful ligation required a 1:4 ratio of insert to vector, using 20ng of purified PCR product and ~80ng of the purified pREC7/*NdeI* vector. The new *Ta*CS construct was transformed into the citrate synthase negative *E.coli* stain, MOB154, using the heat-shock method (see Materials and Methods). 50ng of the ligation mixture was transformed into

100µl of prepared MOB154 competent cells. This transformation resulted in ~120 colonies on one LB-ampicillin plate incubated at 37°C overnight. Eight of these colonies were selected and the plasmid DNA was isolated from 1ml of overnight culture of each colony. Figure 4.7 shows the restriction digests, with *NdeI* and *KpnI*, of the plasmid DNA from the eight selected colonies. Two of the eight selected colonies have been successfully transformed with the correct construct, having bands corresponding to the insert (1.3kb), vector (2.5kb) and the partially digested full construct (3.8kb).

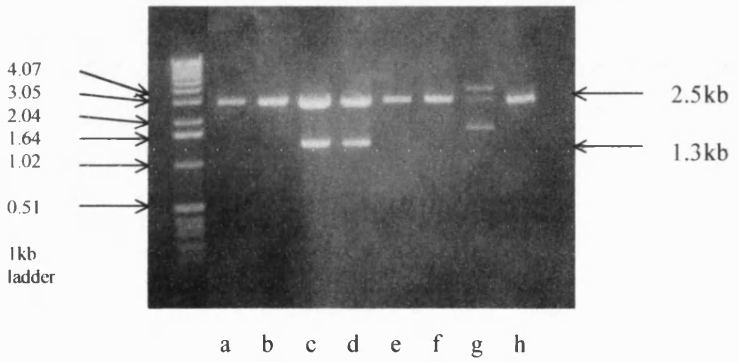


Figure 4.7 **Restriction digest of *Ta*CS wt in pREC7/NdeI**
The figure shows the plasmid DNA prepared from 8 colonies of MOB154 cells transformed with the *Ta*CS wt gene/pREC7/NdeI construct. Lanes a-h show the plasmid digested with NdeI and KpnI restriction enzymes. Lanes c and d have bands that correspond to the singly cut full construct (3.8kb), the pREC7/*NdeI* expression vector (2.5kb) and the *Ta*CS wt gene insert (1.3kb). Lanes a, b, e, f, and h show a single band corresponding to the vector (2.5kb) indicating that the ligation of the insert into the vector was unsuccessful in these cases. Lane g has a band that may correspond to the vector, but has the incorrect insert. The plasmid preperation run in Lane d was selected for purification of *Ta*CS wt.

One colony was selected for the expression and purification of the *Ta*CS wt. The *Ta*CS wt gene from the selected colony was sequenced using the RAM1 primer (forward) and the RAM2 primer (reverse). Each sequence reaction gave a DNA sequence of ~800bp which were compared with the original *Ta*CS wt sequence using the GCG program. Figure 4.8 shows a double digest of the construct from the selected colony along with the single digests to confirm that a *Ta*CS wt construct has been successfully created. Expression of *Ta*CS wt from this construct is discussed in section 4.5.

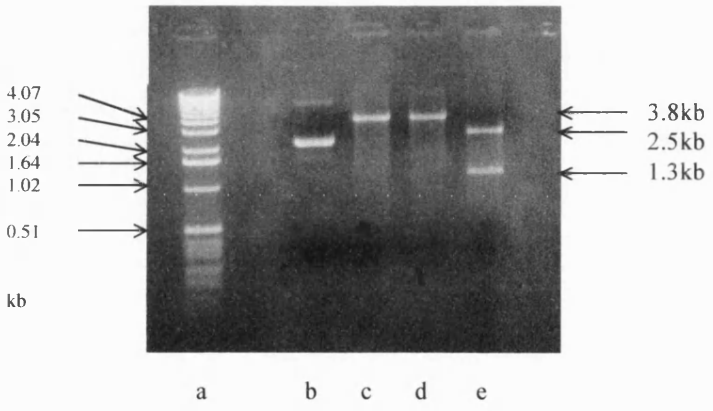


Figure 4.8 **Single and double *NdeI* and *KpnI* restriction digests of pREC7/*NdeI*/*Ta*CS wt construct**
Lane a shows the DNA marker (1kb). Lane b is the undigested vector/insert construct. Lanes c and d show single digests of the construct with *NdeI* and *KpnI* respectively. Lane e is the double restriction digest with *NdeI* and *KpnI*, and shows that it is the correct construct, having bands corresponding to the insert (1.3kb), vector (2.5kb) and the partially digested full construct (3.8kb).

4.4 Site-directed mutagenesis of *Ta*CS wt gene.

The *Ta*CS ionic network mutants were constructed using the nested-PCR (or megaprimer) method outlined in section 4.2.1. The first round PCR was the standard PCR method as described in Materials and Methods. This protocol is outlined below:

50ng template DNA

15pmoles each primer (RAM1 plus required mutation primer)

20mM dNTPs

×10 Thermopol buffer

2U Vent DNA polymerase

The temperature cycles were optimized for the binding of the 1st round PCR primers (each between 26-42 bases in length). The vent polymerase was added to the PCR reaction mix after a 'hot start':

96°C for 5 min

85°C for 3 min

The 1st round DNA fragment was then amplified by 30 cycles of:

96°C for 1.25 min

45°C for 2 min

72°C for 2 min

The PCR was then completed with a 10-minute incubation at 72°C followed by a soak at 4°C.

Figure 4.9 depicts a typical result from a 1st round PCR. As all the mutations are in the same region of the *Ta*CS wt gene (bases 524-614), see Figure 4.4, all the 1st round PCR products are expected to be between 300-400 bases in length. The 1% agarose gel shown in Figure 4.9 shows the 1st round PCR product in the construction of *Ta*CS 4sm (T111E).

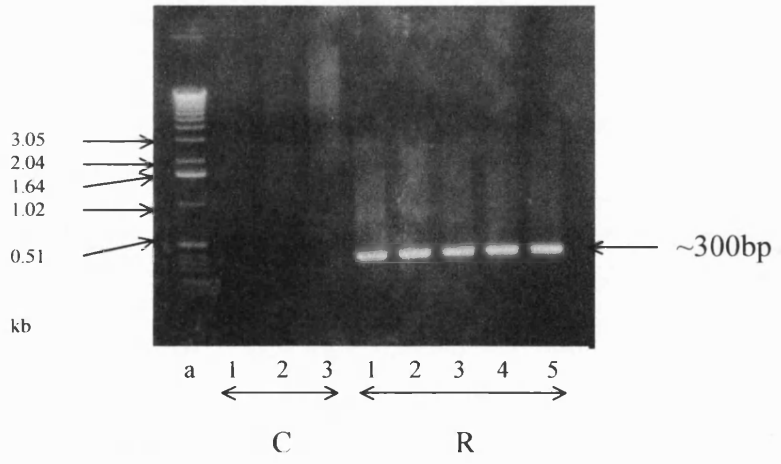


Figure 4.9 **1st round PCR in the construction of *Ta*CS 4sm (T111E).** Lane a shows DNA marker (1kb). Lanes C1, C2 and C3 show the PCR controls: C2 and C3 are single primer controls and C3 shows the double primer control in the absence of template DNA. Lanes R1-5 are multiple 1st round PCR products showing a product of the expected size (~300bp).

As illustrated in Figure 4.3, the 2nd round PCR in the nested-PCR method utilizes the 300-400bp 1st round PCR product as a 'megaprimer'. The conditions for the 2nd round PCR are similar to those of the 1st round PCR, the important changes being the increased time of the primer annealing step and the concentration of the megaprimer.

As the megaprimer is much longer than a standard primer (300-400 bases instead of ~30 bases) then the annealing time must be increased to ensure that the full-length megaprimer binds to the template DNA. The concentration of the megaprimer is also important for this reason. Brons-Poulsen *et al.* (1998) showed that an increased concentration of the megaprimer increases the yield of product. Hence, the annealing time for the 2nd round PCR was increased to 5 min and the concentration of the megaprimer giving optimal product yield was found to be at ~2-2.5 pmoles. The vent polymerase was added to the reaction mix after a 'hot start' as before, then the DNA fragment was amplified by 30 cycles of:

96°C for 1.25 min

50°C for 5 min

72°C for 2 min

The PCR was then completed with a 10-minute incubation at 72°C followed by a soak at 4°C as before.

A typical result for the 2nd round PCR is shown in Figure 4.10. This figure shows the results of the 2nd round PCR in the construction of *Ta*CS 3sm (M100R). For each mutant the 2nd round PCR product is expected to be the full length *Ta*CS gene (1.3kb) with the mutations incorporated.

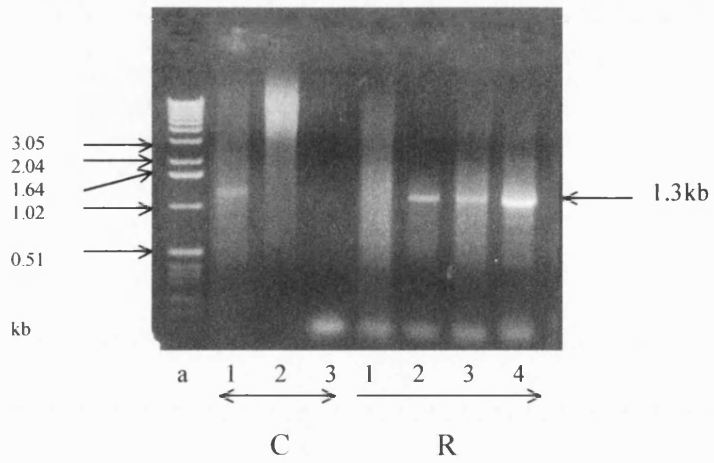


Figure 4.10 **2nd round PCR in the construction of *Ta*CS 3sm (M100R).** Lane a shows the DNA marker (1kb).). Lanes C1, C2 and C3 show the PCR controls: C2 and C3 are single primer controls and C3 shows the double primer control in the absence of template DNA. Lanes R1-4 are 2nd round PCR products showing a product of the expected size (~1.3kb). The arrow indicates PCR reactions with increasing concentrations of 1st round PCR mega-primer

Obviously different primers are required for each different mutation in the 1st round PCR. In each case this gives a 300-400 base fragment that is able to introduce a single mutation into any *Ta*CS template (*Ta*CS wt or a *Ta*CS that already contains a mutation) during the 2nd round PCR. This is useful as it enables many combinations of mutations to be

introduced into the *Ta*CS wt. Table 4.3 shows the primer required for the 1st round PCR (apart from RAM1) and the *Ta*CS template used in the construction of each *Ta*CS mutant.

*Ta*CS name	Mutation	Primer used for 1st round PCR (apart from RAM1)	*Ta*CS template used for both 1st and 2nd round PCRs
*Ta*CS 3sm	M100R	RAM3	*Ta*CS wt
*Ta*CS 4sm	T111E	RAM4	*Ta*CS wt
*Ta*CS 5sm	D94H	RAM5	*Ta*CS wt
*Ta*CS 3/4dm	M100R/T111E	RAM4	*Ta*CS 3sm
*Ta*CS 4/5dm	T111E/D94H	RAM5	*Ta*CS 4sm
*Ta*CS 3/4/5tm	M100R/T111E/D94H	RAM5	*Ta*CS 3/4dm
*Ta*CS 6sm	G/D/I/P	RAM6	*Ta*CS wt
*Ta*CS 6/3dm	M100R/G/D/I/P	RAM6	*Ta*CS 3sm
*Ta*CS 6/5dm	D94H/G/D/I/P	RAM6	*Ta*CS 5sm
*Ta*CS 3/5/6tm	M100R/D94H/G/D/I/P	RAM5	*Ta*CS 6/3dm

Table 4.3 **Summary of the mutagenic primers and template DNA used in the construction of each *Ta*CS ionic network mutant.**

All 2nd round PCR products were ligated into the pREC7/*NdeI* expression vector and transformed into the citrate synthase negative *E.coli* strain MOB154 as for *Ta*CS wt (see Section 4.3). In each case, one transformed colony was selected for the expression of each *Ta*CS ionic network mutant. The *Ta*CS mutants genes were sequenced using primers

RAM1 and RAM2, and all contained only the required mutations. Figure 4.11 shows the double restriction digest, with *NdeI* and *KpnI*, of the plasmid DNA isolated from the selected colony for each *Ta*CS mutant indicating that all colonies have been successfully transformed with the correct construct, having bands corresponding to the insert (1.3kb), vector (2.5kb) and the partially-digested full construct (3.8kb).

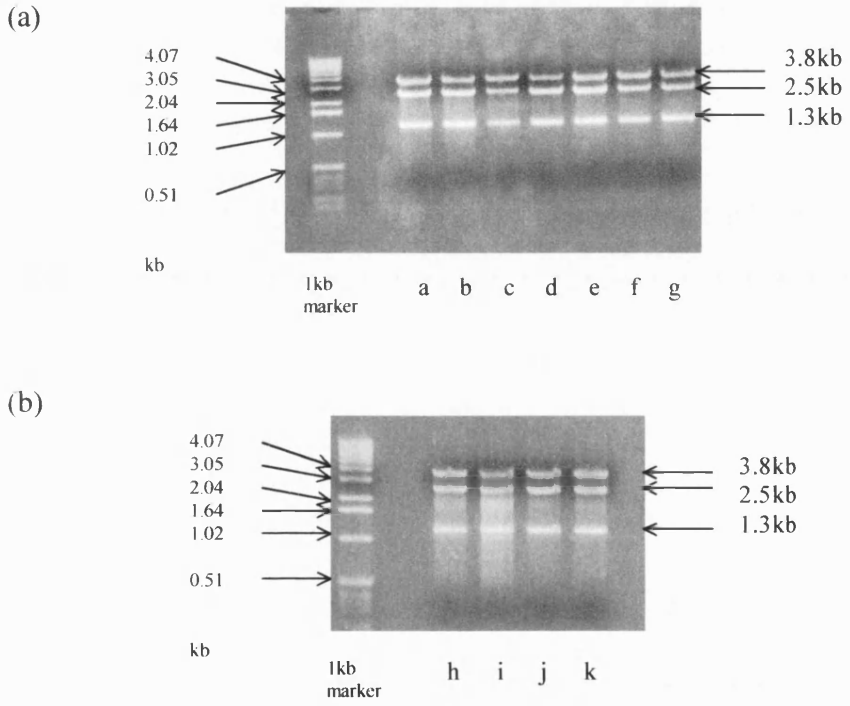


Figure 4.11 **Restriction digests of all *Ta*CS/pREC7/*NdeI* constructs.**
These gels shows that each mutant has the correct construct, having bands corresponding to the insert (1.3kb), vector (2.5kb) and the partially digested full construct (3.8kb). On gel (a) lane a shows TaCS wt, lane b shows *Ta*CS 3sm, lane c shows *Ta*CS 4sm, lane d shows *Ta*CS 5sm, lane e shows *Ta*CS 3/4dm, lane f shows *Ta*CS 4/5dm and lane g show *Ta*CS 3/4/5tm. On gel (b) lane h shows *Ta*CS 6sm, lane i shows *Ta*CS 6/3dm, lane j shows *Ta*CS 6/5dm and lane k shows *Ta*CS 3/5/6tm.

4.5 Purification of wt and mutant *Ta*CSs

The transformed MOB154 cells for *Ta*CS wt and all the *Ta*CS ionic network mutants were grown in a 500ml culture of Terrific Broth (100μg ampicillin/ml) to an optical density at 600nm of 5, then induced with 50μg nalidixic acid/ml. For comparative purposes the citrate synthase from *Ss*CS wt was also expressed in pREC7/*NdeI* in MOB154 cells. In each case, between 5-7g of cells were obtained after centrifugation and resuspended in 2ml of 20mM EPPS (pH 8.0), 2mM EDTA. Cell extracts were prepared by sonication (see methods) and then heat-treated by incubation at 65°C for 15 min.

A 2ml Dye Matrex Red A column was equilibrated with 10×bed volume of 20mM EPPS (pH 8.0), 2mM EDTA, and then the 2ml heat treated samples were applied to the column. Another 10×bed volume of 20mM EPPS (pH 8.0), 2mM EDTA was used to wash the column, after which the citrate synthase was eluted with a solution of 5mM OAA and 1mM CoA. Ten 1ml fractions were collected in each case. Figure 4.12 shows a typical elution of the *Ta*CS wt and *Ta*CS mutants, illustrating that the citrate synthase is present in the first five fractions collected, which was the case for each citrate synthase purified. These data are consistent with SDS-PAGE analysis of the first five fractions collected for each citrate synthase (Figure 4.13). Purification tables for all the citrate synthases purified are shown in the Appendix; 0.5-5mg of protein were obtained in each case and each protein was purified between 5-20 fold.

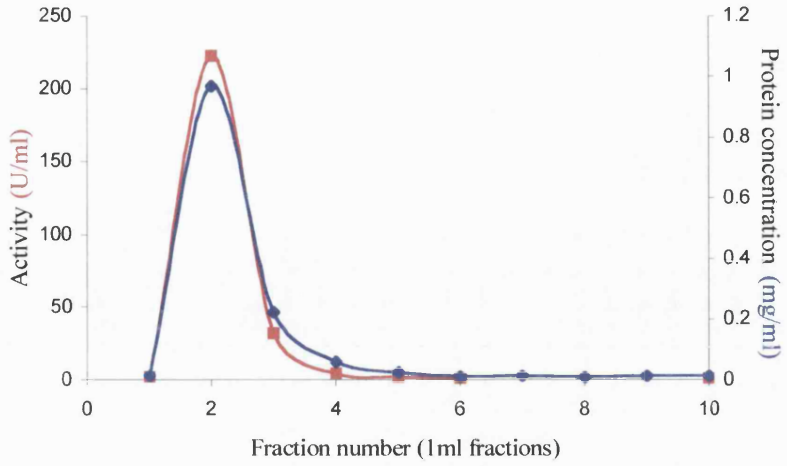


Figure 4.12 ***Ta*CS wt enzyme elution from Dye Matrex Red A column.**
Graph shows the activity and protein concentration of each of the ten fractions collected after elution of the *Ta*CS wt from the Dye Matrix Red A column. The activity (U/ml) is shown in red and the protein concentration (mg/ml) is shown in blue.

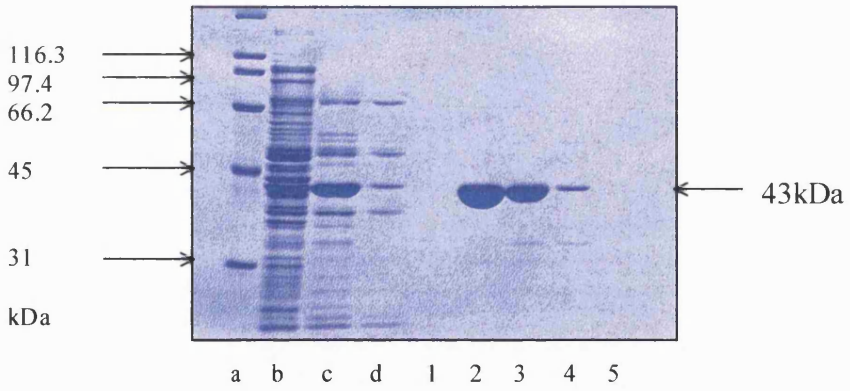


Figure 4.13 **SDS-PAGE of *Ta*CS wt purification.**
SDS-PAGE shows the protein content of five purified fractions of *Ta*CS wt eluted from the Dye Matrex Red A column. Lanes 1-5 correspond to the first five fractions shown in Figure 4.12. Lane a is a protein marker (broad range). Lane b corresponds to the crude extract, lane c to the heat-treated extract (65°C for 15 min) and lane d to the flowthrough from the column collected before the ligand-induced elution of the protein.

Figure 4.14 shows that all the purified *Ta*CS samples were homogeneous on SDS-PAGE and all the enzymes had Mr values of ~43, which is consistent with previous data [Sutherland *et al.*, 1991]. The *Ss*CS wt can be seen as a smaller band (~41-42 kDa) which is consistent with previous purifications of this enzyme [Smith *et al.*, 1987]. Unfortunately, the *Ss*CS wt showed the presence of a contaminating band with a smaller molecular weight. Due to time constraints it was impossible to rectify this problem, so the *Ss*CS wt was used as shown in Figure 4.14. The contamination present in the *Ss*CS wt sample will affect the specific activity of this enzyme and may affect the overall thermostability depending on the nature of the contaminating protein. This problem is considered when characterizing the *Ss*CS wt in Chapter 5, and data obtained are compared with previous data obtained for this enzyme.

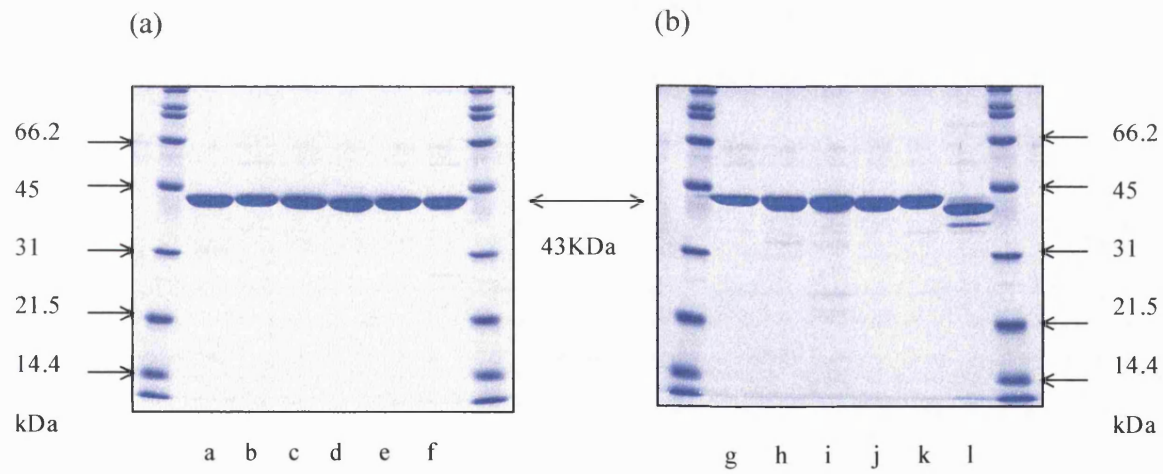


Figure 4.14 **SDS-PAGE of all purified *Ta*CS enzymes.**
Gel shows each *Ta*CS enzyme after elution from Dye Matrex Red A column. Each lane contains ~4µg protein. On gel (a) lane a- *Ta*CS wt, lane b- *Ta*CS 3sm, lane c- *Ta*CS 4sm, lane d- *Ta*CS 5sm, lane e- *Ta*CS 3/4dm and lane f- *Ta*CS 4/5dm. On gel (b) lane g- *Ta*CS 3/4/5tm, lane h- *Ta*CS 6sm, lane i- *Ta*CS 6/3dm, lane j- *Ta*CS 6/5dm and lane k- *Ta*CS 3/5/6tm and lane l- *Ss*CS wt. Samples are run against broad-range protein markers.

4.6 Discussion

The nested-PCR (mega primer) method of introducing mutations into the *Ta*CS wt proved to be successful. The two-step nature of this method enabled the production of many combinations of ionic network mutants. This is advantageous when introducing a number of charged residues into a localized area of a protein as discussed by Lebbink *et al.* (1999). This group aimed to introduce the equivalent 18-residue ionic network present at the subunit interface of glutamate dehydrogenase (GDH) from the hyperthermophilic *P.furiosus* into the GDH from the less hyperthermophilic *Thermotoga maritima*. This involved introducing four new charged amino acids into the subunit interface of *T.maritima* GDH. Interestingly, the data showed that successive introduction of the amino acids substituted to form the network did not result in incremental stabilization of the GDH enzyme. However, several combinations of individual destabilizing mutations resulted in restored activity and stability. This suggests that there is high cooperativity between the amino acid members of the ionic network. Hence, the ionic network was introduced into the *Ta*CS wt in multiple combinations of the four separate mutations, M100R (RAM3), T111E (RAM4), D94H (RAM5) and G/D/I/P (RAM6).

The *Ta*CS wt and the *Ta*CS ionic network mutants were successfully cloned into the pREC7/NdeI expression vector (Figure 4.11) and all were sequenced to confirm that only the desired mutations were present in each gene. The enzymes were expressed in MOB154 citrate synthase negative *E.coli* cells and successfully purified on a small scale Dye Matrex Red A column (2ml). Purification tables (appendix) show that 0.5-5mg of protein were

obtained. As these tables show, in some cases the yield of protein obtained is not exceptional i.e. only 5% recovery for the purification of *Ta*CS 6/3dm. It was only necessary to purify enough protein for complete characterization of each enzyme, and the quantities of protein obtained were sufficient for this purpose. Figure 4.14 shows that all of the *Ta*CSs appear, on SDS-PAGE, to be purified to a similar degree of homogeneity. All recombinant enzymes are shown to have a molecular weight of ~43, which is consistent with previous purifications of both native [Smith *et al.*, 1987] and recombinant [Sutherland *et al.*, 1991] *T.acidophilum* citrate synthase.

Preliminary thermal inactivations of the *Ta*CS ionic network mutants suggested that at least one of the *Ta*CS mutants might show an increase in thermostability. This increase in thermal stability was not equal to the thermostability of *Pf*CS wt, but more comparable to the thermostability of the citrate synthase from *Ss*CS. For this reason it was decided to purify the recombinant *Ss*CS wt (supplied by Carl Thompson, University of Bath), for comparative purposes. Previous purifications of the *Ss*CS wt show that the native enzyme has a molecular weight of ~41kDa [Smith *et al.*, 1987] and the recombinant has a Mr of ~42.5kDa [Connaris *et al.*, 1998] which correspond to the major band present in Figure 4.14. Unfortunately, there is also a contaminating band present that is smaller and may be a degradation product. This is important, as any contamination will effect the calculation of the specific activity of this enzyme. Initial specific activity data presented in the purification tables (appendix) show that the activity of this *Ss*CS wt is comparable with the specific activities of the *Ta*CS wt and some of the mutants (~100 U/mg). Previous specific activity data show that the specific activity of *Ss*CS wt is usually ~30-55% higher than

*Ta*CS wt at 55°C [Smith *et al.*, 1987, James *et al.*, 1994]. The purification tables (appendix) show that the specific activity at 55°C of SsCS wt is 17.5% higher than that of *Ta*CS wt. This may illustrate the effect of the contaminating band of *Ss*CS wt, but further characterization of all enzymes must be carried out before any conclusions may be drawn.

The purpose of purifying the citrate synthase from *Sulfolobus solfataricus* was to compare the thermal stability of this enzyme to the thermostability of *Ta*CS wt and the *Ta*CS ionic network mutants. If the contamination of the *Ss*CS wt does not affect the thermostability of this enzyme then comparisons may still be made. Kinetic and thermal characterization of all citrate synthases discussed in this chapter are discussed in Chapter 5.

CHAPTER 5

Kinetic and thermal characterization of the *Thermoplasma acidophilum* citrate synthase ionic network mutants.

5.1 Introduction

As explained in the Chapter 4, ten mutants of the *Thermoplasma acidophilum* citrate synthase were created in an attempt to introduce an ionic network at the subunit interface that would mirror the ionic network found in *Pyrococcus furiosus* citrate synthase. This chapter describes and discusses the kinetic and thermal characterization of these mutant enzymes with relation to the *T.acidophilum* wild-type citrate synthase (*Ta*CS wt).

Single and multiple combinations of four mutations were introduced into the *Ta*CS wt gene using a nested (or megaprimer) PCR method (Section 4.2). Multiple combinations were used as it has been previously shown in a similar study that some single mutations may result in a decrease towards thermal inactivation, but a combination of destabilizing single mutations can restore stability [Lebbink *et al.*, 1999]. Table 4.2 (Chapter 4) summarises the *Ta*CS mutants created, indicating each mutation that the enzyme contains.

5.2 Kinetic parameters of *Ta*CS wt and the *Ta*CS ionic network mutants.

For each *Ta*CS, the K_m values for oxaloacetate (OAA) and acetyl-CoA (AcCoA) and the V_{max} were determined. As for previous kinetic data (Section 3.5), the K_m for each enzyme was determined by adding a fixed saturating concentration of one substrate (10-15 × K_m) and varying the concentration of the other (0.5, 0.7, 1, 5 and 10 × K_m). The concentrations of substrate were based on previous kinetic parameters determined for *Ta*CS wt [Arnott *et al.*, 2000]. The data were analysed using ENZPACK, and the K_m and V_{max} values were obtained using the Direct Linear plot [Eisenthal and Cornish-Bowden, 1974]. As before, the data obtained follow a Michaelis-Menten relationship and the straight line Hanes-Wolfe plots (as illustrated in Figures 3.8 and 3.9, Chapter 3). Table 5.1 summarises the kinetic data for *Ta*CS wt and the ionic network mutants. The table also includes kinetic data obtained for the *S.solfataricus* citrate synthase (*Ss*CS wt) purified as described in the previous chapter (Section 4.5).

Table 5.1 shows that, in most cases, the substitution of residues at the subunit interface of *Ta*CS results in an increase of the K_m for AcCoA. Only *Ta*CS 3sm, *Ta*CS 4sm, *Ta*CS 3/4dm and *Ta*CS 6sm showed an increase of less than 3-fold K_m(AcCoA) compared to TaCS wt. Every other *Ta*CS mutant showed a K_m(AcCoA) increase of more than 3-fold compared to *Ta*CS wt, and some mutants had up to a 10-fold increase in this value (TaCS 4/5dm).

Citrate synthase	K_m AcCoA (μM)	K_m OAA (μM)	V_{max} (μmol/min/mg)	k_{cat} (s^{-1})
*Ta*CS wt	3	6	77	55
*Ta*CS 3sm	7	12	146	105
*Ta*CS 4sm	5	13	96	69
*Ta*CS 5sm	18	13	80	57
*Ta*CS 3/4dm	8	9	143	102
*Ta*CS 4/5dm	29	6	73	52
*Ta*CS 3/4/5tm	15	2	41	29
*Ta*CS 6sm	4	4	74	53
*Ta*CS 6/3dm	11	17	65	47
*Ta*CS 6/5dm	21	10	86	62
*Ta*CS 3/5/6tm	13	11	51	37
*Ss*CS wt	7	48	127	91

Table 5.1 **Kinetic parameters determined for *Ta*CS wt, *Ss*CS wt and *Ta*CS ionic network mutants.**
The K_m values are determined by fixing the concentration of one substrate (>20×K_m), and varying the concentration of the other. The V_{max} were calculated using the data from the Direct Linear plot in ENZPACK and the determination of protein concentration (mg). Standard errors were determined from the standard deviations (68% confidence limits) given from the Direct Linear Plot analyses, and were <10% on all values of K_m and V_{max}. The *Ta*CS mutants are described in Section 4.1. The k_{cat} values were calculated using a Mr value of 43 in each case.

The K_m(OAA) values for the *Ta*CS ionic network mutants seem to be less affected than the K_m(AcCoA) values by the amino acid substitutions at the subunit interface. There is a small increase in K_m(OAA) values for most of the *Ta*CS mutants, but this does not exceed 2.5-fold of the K_m(OAA) for *Ta*CS wt. *Ta*CS 3/4/5tm and *Ta*CS 6sm show a slight decrease in K_m(OAA) compared to the *Ta*CS wt, and *Ta*CS 4/5dm shows little change.

The V_{max} for *Ta*CS wt is 77 μmol/min/mg and, for comparative purposes, any *Ta*CS mutant V_{max} value between 70 and 80 μmol/min/mg is considered relatively unchanged. An increase in V_{max} is any value greater than 80 μmol/min/mg, and a decrease is any V_{max} value below 70 μmol/min/mg. Using these criteria, the majority of the *Ta*CS mutants have relatively unchanged V_{max} values compared to the *Ta*CS wt. Both of the *Ta*CS mutants containing the full ionic networks (*Ta*CS 3/4/5tm and *Ta*CS 3/5/6tm) show a decrease in V_{max}, along with the *Ta*CS 6/3dm mutant. *Ta*CS 3sm and *Ta*CS 3/4sm show an increase in V_{max} and k_{cat} values.

5.3 Thermal inactivation of *Ta*CSs.

The *Ta*CS wt and the *Ta*CS ionic network mutants were analysed for their thermostability properties in terms of their resistance to irreversible denaturation. Each enzyme was incubated at a chosen temperature (65-85°C) and, at defined time intervals, samples were snap-cooled in an ice-water bath and then assayed for enzyme activity at 55°C under standard assay conditions (pH 7.0). To investigate the effect of the presence of substrates on the thermostability of these citrate synthases, the thermal inactivations were carried out in both the absence and presence of citrate synthase substrates.

5.3.1 Thermal inactivation in the absence of substrates.

Figures 5.1, 5.2, 5.3, 5.4 and 5.5 show the thermal inactivations of the *Ta*CSs at 65°C, 70°C, 75°C, 80°C and 85°C respectively, in the absence of substrates.

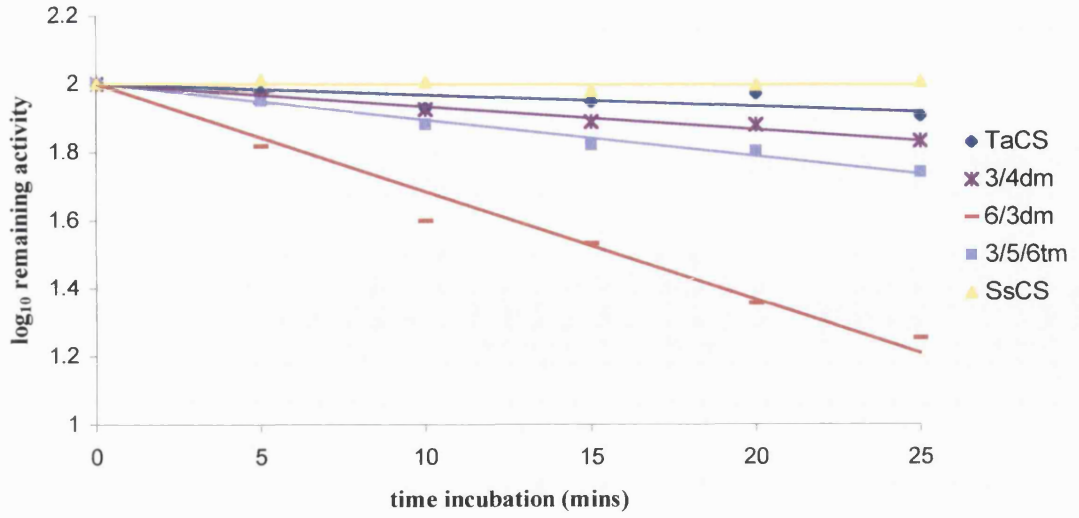


Figure 5.1 **Thermal inactivation of citrate synthases at 65°C, in the absence of substrates.**
This figure shows only the data obtained for *Ta*CS mutants that showed a decrease in thermostability compared to the *Ta*CS wt and *Ss*CS wt. All *Ta*CS mutants that are not shown (*Ta*CS 3sm, *Ta*CS 4sm, *Ta*CS 5sm, *Ta*CS 4/5dm, *Ta*CS 3/4/5tm, *Ta*CS 6sm and *Ta*CS 6/5dm) showed thermostabilities similar to that of *Ta*CS wt at this temperature.

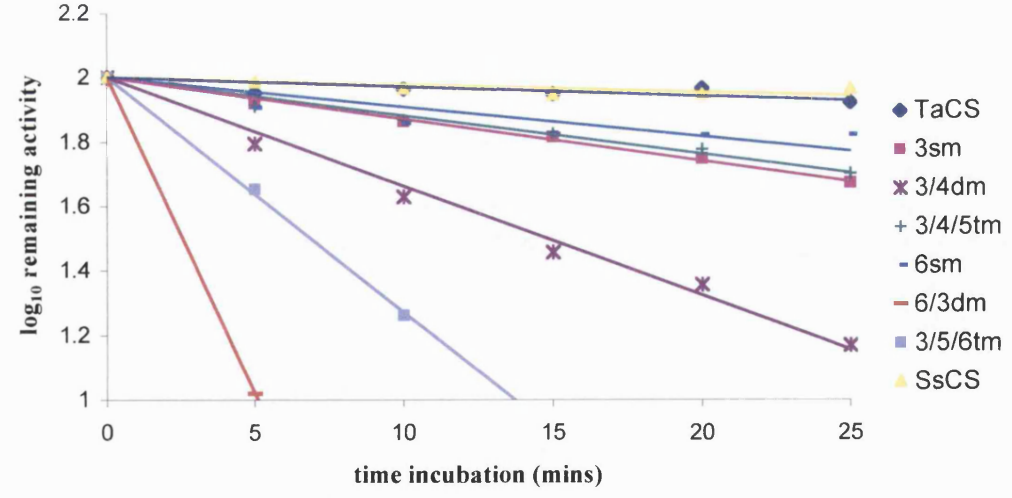


Figure 5.2 **Thermal inactivation of citrate synthases at 70°C, in the absence of substrates.**
Only the data obtained for *Ta*CS mutants showing a decrease in thermostability compared to *Ta*CS wt at this temperature are shown in the figure. The *Ta*CS mutants not shown (*Ta*CS 4sm, *Ta*CS 5sm, TaCS 4/5dm and *Ta*CS 6/5dm) showed similar thermostabilities to that of the *Ta*CS wt at 70°C.

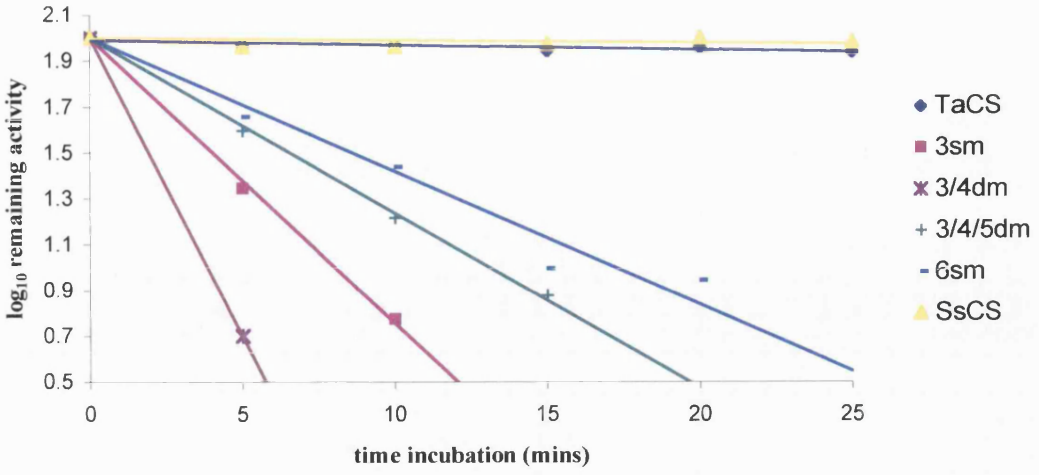


Figure 5.3 **Thermal inactivation of citrate synthases at 75°C, in the absence of substrates.**
*Ta*CS 6/3dm and *Ta*CS 3/5/6tm are not shown on this figure as they had no remaining activity after incubation at 75°C in the timescale indicated. *Ta*CS 4sm, *Ta*CS 5sm, *Ta*CS 4/5dm and *Ta*CS 6/5dm are not shown as they displayed similar thermal stabilities to the *Ta*CS wt at this temperature.

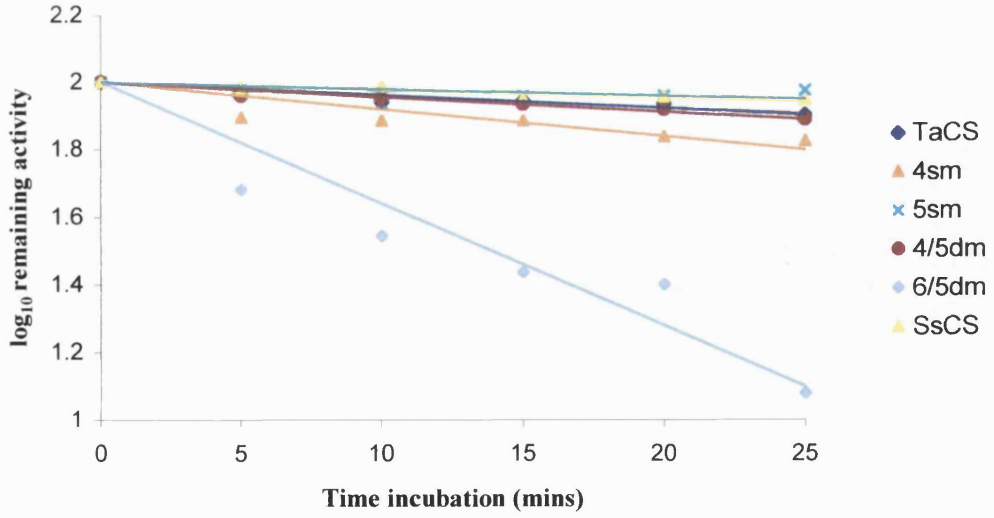


Figure 5.4 **Thermal inactivation of citrate synthases at 80°C, in the absence of substrates.**
All *Ta*CS ionic network mutants not shown on this figure (*Ta*CS 3sm, *Ta*CS 3/4dm, *Ta*CS 3/4/5tm, *Ta*CS 6sm, *Ta*CS 6/3dm and *Ta*CS 3/5/6tm) do not display any remaining activity after incubation at 80°C in the timescale indicated.

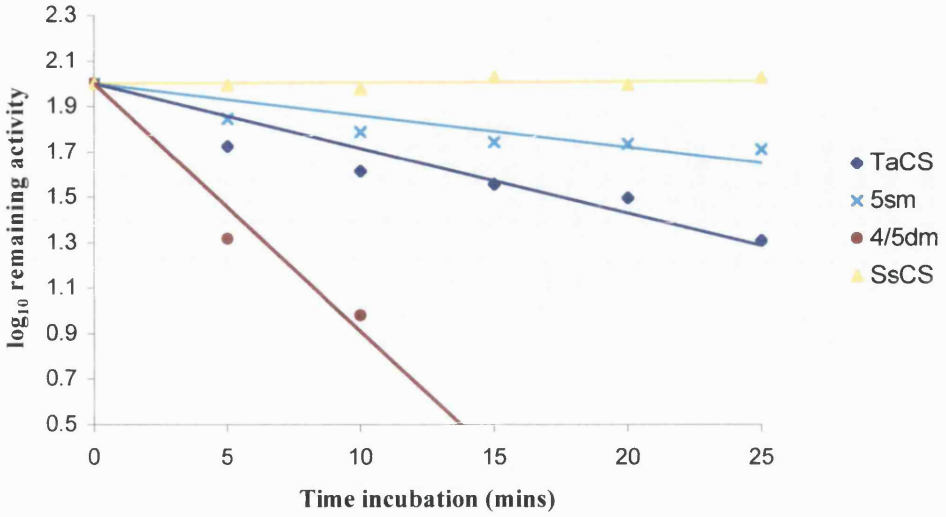


Figure 5.5 **Thermal inactivation of citrate synthases at 85°C, in the absence of substrates.**
All *Ta*CS ionic network mutants not shown in this figure (*Ta*CS 3sm, *Ta*CS 4sm, *Ta*CS 3/4dm, *Ta*CS 3/4/5tm, *Ta*CS 6sm, *Ta*CS 6/3dm, *Ta*CS 6/5dm and *Ta*CS 3/5/6tm) do not display any remaining activity after incubation at 85°C in the timescale indicated.

For illustrative purposes the data obtained for some *Ta*CS mutants are omitted from certain figures depending on whether they retain no activity at that incubation temperature or show no great difference in thermal inactivation from the *Ta*CS wt; this is stated in the figure legend of each figure. Table 5.2 displays all the first-order rate constants (k_{inact}) and the $t_{½}$ (half-life) values obtained for these thermal inactivations of *Ta*CS wt and the *Ta*CS ionic network mutants.

citrate synthase	65°C		70°C		75°C		80°C		85°C	
	k_{inact} (min^{-1})	$t_{½}$ (min)	k_{inact} (min^{-1})	$t_{½}$ (min)	k_{inact} (min^{-1})	$t_{½}$ (min)	k_{inact} (min^{-1})	$t_{½}$ (min)	k_{inact} (min^{-1})	$t_{½}$ (min)
*Ta*CS wt	0.0074	94.4	0.0067	102.4	0.0062	109.9	0.0088	78.6	0.0659	10.5
*Ta*CS 3sm	0.0090	76.6	0.0297	23.2	0.2867	2.4	×	×	×	×
*Ta*CS 4sm	0.0023	304.7	-	-	0.0044	156	0.0184	37.5	×	×
*Ta*CS 5sm	0.0023	292.3	-	-	0.0039	174	0.0046	152.8	0.0322	21.5
*Ta*CS 3/4dm	0.0138	44.3	0.0778	8.9	0.5997	1.2	×	×	×	×
*Ta*CS 4/5dm	0.0018	387.9	-	-	0.0078	88.5	0.0101	68.9	0.2512	2.8
*Ta*CS 3/4/5tm	0.0064	107.5	0.0274	25.3	0.1748	4	×	×	×	×
*Ta*CS 6sm	0.0064	109.1	0.0212	32.8	0.1338	5.2	×	×	×	×
*Ta*CS 6/3dm	0.0645	10.8	0.4528	1.5	×	×	×	×	×	×
*Ta*CS 6/5dm	0.0009	720.2	-	-	0.0074	93.8	0.0831	8.3	×	×
*Ta*CS 3/5/6tm	0.0219	31.8	1.6803	4.1	×	×	×	×	×	×
*Ss*CS wt	0.0002	2150	0.0053	132	0.0025	276.2	0.0053	132	0.0014	524.4

Table 5.2 **Thermal inactivation parameters, k_{inact} and $t_{½}$, of *Ta*CS ionic network mutants compared to *Ta*CS wt and *Ss*CS wt over a range of temperatures.**
All data displayed in this table are calculated from the thermal inactivations, in the absence of substrates, shown in Figures 5.1-5.5 (including some data not shown in these figures for illustrative purposes). The - symbol indicates where there are no data available, but there is no great change in thermostability compared to the *Ta*CS wt after incubation at the indicated temperature. The × symbol indicates where the CS enzymes no longer show residual activity after incubation at the relevant temperature. All k_{inact} data have standard errors <10%, not shown.

Interestingly, most of the *Ta*CS ionic network mutants seem to have a decrease in k_{inact} value and an increase in $t_{½}$ value compared to the *Ta*CS wt at 65°C (apart from *Ta*CS 3sm, *Ta*CS 3/4dm, *Ta*CS 6/3dm and *Ta*CS 3/5/6tm). The ionic network mutants that show this increased thermostability at 65°C all contain the M100R mutation (RAM3), illustrated in Figure 5.1. *Ta*CS 3/4/5tm is the only mutant that contains the M100R mutation that has a reduction in k_{inact} and an increase in $t_{½}$, but these are only marginal changes compared to the *Ta*CS wt at 65°C (not shown on Figure 5.1).

It is clear from Figures 5.1-5.5 and Table 5.2 that the rates of thermal inactivation are different for each different combination of mutation although some trends consistent with a specific mutation can be identified. Figure 5.2 shows that a thermal inactivation at 70°C results in the loss of activity of all mutants containing either the M100R (RAM3) or G/D/I/P loop insertion (RAM6) mutations, apart from *Ta*CS 6/5dm. All the *Ta*CS ionic network mutants containing the M100R (RAM3) mutation are inactive or have a half-life <4 minutes with a thermal inactivation at 75°C (Figure 5.3).

The only *Ta*CS ionic network mutants that show residual activity after thermal inactivation at 80°C are those with single or combination mutations of T111E (RAM4) and D94H (RAM5), (Figure 5.4). The *Ta*CS 6/5dm mutant still retains some activity after incubation at 80°C but is greatly reduced compared to the *Ta*CS wt, a ten-fold increase in k_{inact} being found at this temperature (Table 5.2). Inactivation at 85°C (Figure 5.5) suggests that the only *Ta*CS ionic network mutant to show any increase in thermal stability, in the absence of substrates, is the *Ta*CS 5sm (single D94H mutation) that has a two-fold decrease in k_{inact} at this temperature.

Table 5.3 shows the $\Delta\Delta G^*_{inact}$ (the change in activation energy of thermal inactivation, i.e. $\Delta G^*_{wt} - \Delta G^*_{mutant}$) calculated from the first order rate constants for thermal inactivation (k_{inact}) using the Arrhenius equation as rearranged in Chapter 3 (equation 3). The negative values of $\Delta\Delta G^*_{inact}$ indicate cases where the stability of the *Ta*CS mutant is reduced compared to the *Ta*CS wt, and positive values indicate an increase in stability compared to the *Ta*CS wt at each temperature. These data show that an inactivation at 65°C results in an activation energy increase of up to 5kJ/mol for some *Ta*CS mutants compared to the *Ta*CS wt at that temperature. The table shows that as

the temperature of the thermal inactivation is increased the $\Delta\Delta G^*_{inact}$ values for most of the *Ta*CS ionic network mutants decrease. It can be seen from the data that once the difference in activation energy decreases to ~-10-15kJ/mol the $\Delta\Delta G^*_{inact}$ value does not decrease further but the enzyme becomes irreversibly denatured. Table 5.3 shows that there is a 2.1kJ/mol increase in activation energy of the *Ta*CS 5sm mutant compared to the *Ta*CS wt at inactivation at 85°C. This increase is five-fold less than the $\Delta\Delta G^*_{inact}$ increase of *Ss*CS wt compared to the *Ta*CS wt.

Citrate synthase	$\Delta\Delta G^*_{inact}$ of *Ta*CS mutant compared to *Ta*CS wt at each temperature (kJ/mol)				
	65°C	**70°C**	**75°C**	**80°C**	**85°C**
*Ta*CS 3sm	-0.6	-4.2	-11.1	×	×
*Ta*CS 4sm	+3.3	-	+1.0	-2.2	×
*Ta*CS 5sm	+3.3	-	+1.3	+1.9	+2.1
*Ta*CS 3/4dm	-1.8	-7.0	-13.2	×	×
*Ta*CS 4/5dm	+3.8	-	-0.7	-0.4	-4.0
*Ta*CS 3/4/5tm	+0.4	-4.0	-9.7	×	×
TaCS 6sm	+0.4	-3.3	-8.9	×	×
*Ta*CS 6/3dm	-6.1	-12.0	×	×	×
*Ta*CS 6/5dm	+5.9	-	-0.5	-6.6	×
*Ta*CS 3/5/6tm	-3.0	-15.8	×	×	×
*Ss*CS wt	+10.1	+0.7	+2.6	+1.5	+11.5

Table 5.3 **Change in activation energies of *Ta*CS ionic network mutants compared to *Ta*CS wt at various temperatures.**

$\Delta\Delta G^*_{inact}$ values are calculated from k_{inact} data displayed in Table 5.2 using equation 3 (Section 3.4). Values shown are the difference in activation energy ($\Delta\Delta G^*_{inact}$) of the TaCS ionic network mutants compared to *Ta*CS wt. A positive value indicates an increase in thermal stability compared to the *Ta*CS wt, and a negative value indicates a decrease. The - symbol indicates where there are no data available as there is no great difference in thermal stability compared to the *Ta*CS wt in that case. All data shown have standard errors <10%.

5.3.2 Thermal inactivation in the presence of substrates.

Previous thermal inactivation experiments on site-directed mutants of *Pf*CS wt (Section 3.6, Chapter 3) show that the activation energy of thermally induced unfolding of the mutants is increased by the presence of substrates, a 1-3kJ/mol increase in $\Delta\Delta G^*_{inact}$ in the presence of substrates. Also, temperature optima (T_{opt}) determinations and kinetic analyses are necessarily carried out in the presence of both citrate synthase substrates, OAA and AcCoA. For these reasons the thermal inactivations of *Ta*CS wt and the *Ta*CS ionic network mutants were repeated but this time with the addition of OAA and CoA during the thermal incubation. As before, one of the citrate synthase products, CoA, was used instead of AcCoA to prevent any reaction occurring during the incubation period.

Interestingly, it was found that incubation at 65°C and 70°C in the presence of substrates resulted in no loss of catalytic activity over a period of 30 min for any of the citrate synthase enzymes: *Ta*CS wt, *Ta*CS mutants or *Ss*CS wt (data not shown). Figure 5.6 shows the thermal inactivation at 75°C in the presence of substrates (*Ta*CS mutants that were shown to be stable at ≥80°C, in the absence of substrates, are not shown in this figure).

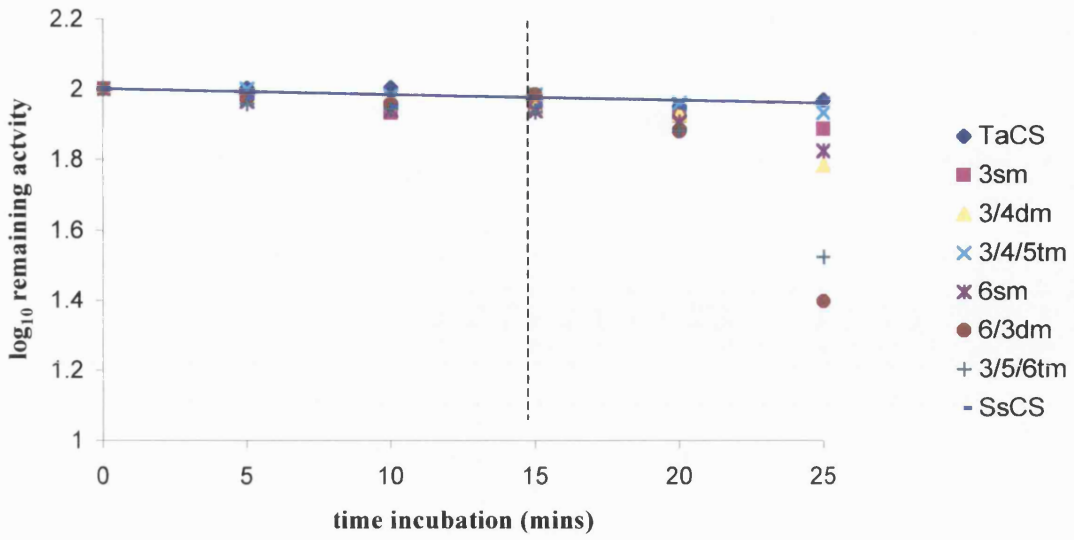


Figure 5.6 **Thermal inactivation of citrate synthases at 75°C, in the presence of substrates.**
Enzymes are incubated at 75°C for up to 25 minutes in the presence of 18mM OAA and 0.2mM CoA. Data for *Ta*CS 4sm, *Ta*CS 5sm, *Ta*CS 4/5dm and *Ta*CS 6/5dm are not shown as they are shown to be stable after incubation at 80°C in Figure 5.4. The dashed line indicates the time until which there appears to be no loss in activity.

Figure 5.6 shows that, for a minimum of 15 minutes, all the *Ta*CS ionic mutants show similar stability to that of the *Ta*CS wt. After 15 minutes of incubation at 75°C (in the presence of substrates) the *Ta*CS mutants that showed a loss of residual activity at this temperature in the absence of substrates (Figure 5.3) rapidly begin to lose residual activity. Similar data were obtained with incubations at 80°C and 85°C although the citrate synthases showed no irreversible denaturation for only 10 and 5 minutes, respectively. OAA is a fairly thermolabile substrate, suggesting that it may be the loss of substrate at these temperatures that results in the inactivation of the enzymes after certain time periods. Therefore, the thermal decomposition of both OAA and CoA at 80°C and 85°C was investigated (Figure 5.7).

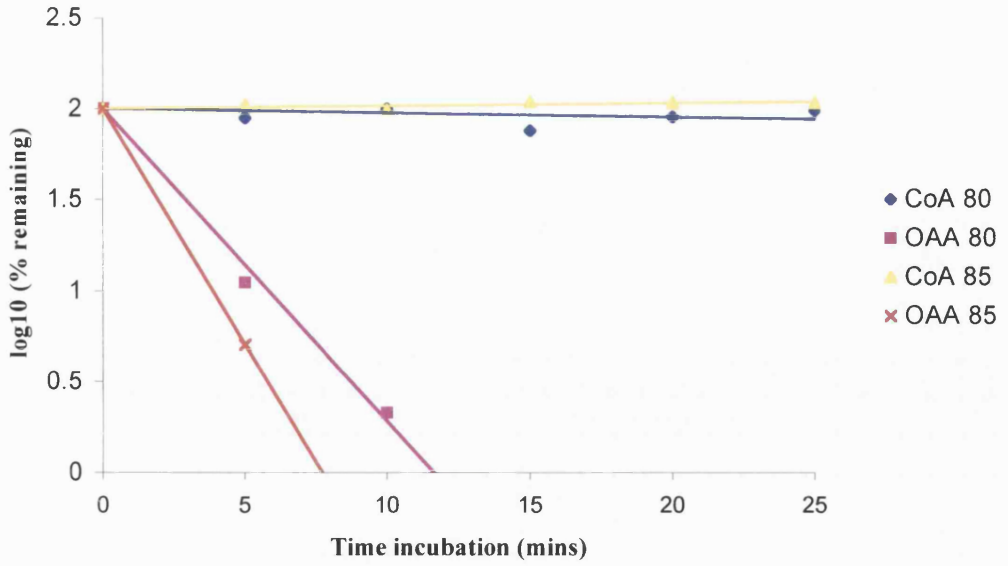


Figure 5.7 **Thermal decomposition of oxaloacetate and CoA at 80°C and 85°C.** 18mM OAA and 0.2mM CoA were incubated at 80 and 85°C and samples were taken out and cooled in an ice-water bath at 5-minute intervals. CoA concentrations were determined by titration with DTNB. Concentrations of OAA were determined in the citrate synthase reaction in the presence of excess AcCoA.

Figure 5.7 shows that there is very little thermal degradation of CoA after 25 minutes incubation at both 80°C and 85°C. Oxaloacetic acid is shown to be much more thermolabile than CoA at these temperatures, with an approximately 115-fold increase rate constant (k_{inact}) at both temperatures. To correlate the thermal degradation of substrate with the thermal inactivations of *Ta*CS in the presence of substrates, it was calculated that after 10 minutes at 80°C, and 5 minutes at 85°C, only 2% and 5% of the original OAA concentrations were present, respectively (18mM initial concentration). This meant that after 10 minutes at 80°C and 5 minutes at 85°C the *Ta*CS enzymes were <95% saturated., and thus susceptible to thermal inactivation.

As there would be a sufficient OAA concentration to give >95% saturation of citrate synthase at 85°C for 5 minutes, the thermal inactivation was repeated in the presence of OAA and CoA for this short time period only (Figure 5.8). The only mutant enzyme

tested in this experiment was *Ta*CS 5sm which had shown increased stability over the *Ta*CS wt enzyme in the absence of substrates (Figure 5.5).

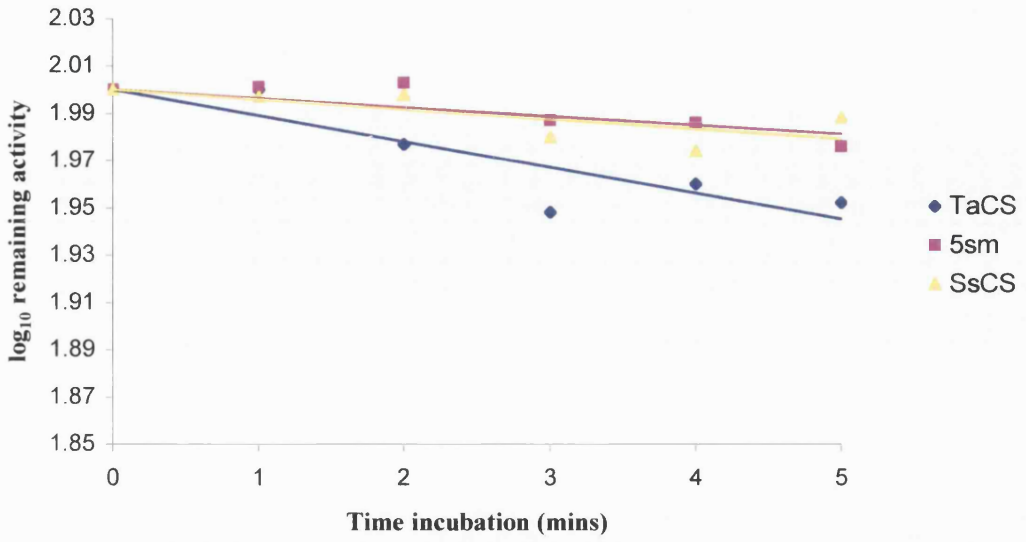


Figure 5.8 **Thermal inactivation of *Ta*CS wt, *Ss*CS wt and the *Ta*CS 5sm ionic network mutant at 85°C, in the presence of substrates.**
The citrate synthase enzymes were incubated with 18mM OAA and 0.2mM CoA. These concentrations are sufficient for the enzymes to be saturated completely for the 5-minute incubation period. All other *Ta*CS ionic mutants are not included in this figure as they showed a decrease in thermal stability at 85°C in the absence of substrates (Figure 5.5).

In the presence of substrates, there is little difference between the thermal stability of *Ta*CS 5sm and that of the *Ss*CS wt (Figure 5.8 and Table 5.4). With respect to the *Ta*CS wt, the k_{inact} values for both *Ta*CS 5sm and *Ss*CS wt, which are both very similar, are approximately 3-fold greater. Table 5.5 shows the $\Delta\Delta G^*_{inact}$ values calculated from Figures 5.5 and 5.8. These data show that both the *Ta*CS wt and the *Ta*CS 5sm ionic network mutant are stabilized by the presence of substrates by ~3-4kJ/mol. Unusually, the thermal stability of the *Ss*CS wt has decreased in the presence of substrates i.e. the activation energy has been reduced by ~6 kJ/mol.

Citrate synthase	k_{inact} (min^{-1})	$t_{½}$ (min)	$\Delta\Delta G^*_{inact}$ compared to *Ta*CS wt (kJ/mol)
*Ta*CS wt	0.0253 ($\pm 3.2\times10^{-3}$)	27.5	-
*Ta*CS 5sm	0.0088 ($\pm 1.8\times10^{-3}$)	79.6	+3.1
*Ss*CS wt	0.0097 ($\pm 2.3\times10^{-3}$)	71.7	+2.9

Table 5.4 **k_{inact} and $t_{½}$ values calculated for the thermal inactivation of *Ta*CS wt, *Ss*CS wt and the *Ta*CS 5sm ionic network mutant at 85°C in the presence of substrates.**
The rates of inactivation (k_{inact}) and half-lives ($t_{½}$) of the citrate synthase enzymes are calculated from data reported in Figure 5.8.

Citrate synthase	$\Delta\Delta G^*_{inact}$ compared to *Ta*CS wt (85°C in presence of substrates) (kJ/mol)	$\Delta\Delta G^*_{inact}$ compared to *Ta*CS wt (85°C in absence of substrates) (kJ/mol)	$\Delta\Delta G^*_{inact}$ of CS in presence of substrates compared to absence of substrates. (kJ/mol)
*Ta*CS wt	-	-	+2.8
*Ta*CS 5sm	+3.1	+2.1	+3.9
*Ss*CS wt	+2.9	+11.5	-5.8

Table 5.5 **$\Delta\Delta G^*_{inact}$ data calculated from rates of thermal inactivation at 85°C in both the absence and presence of substrates.**
The $\Delta\Delta G^*_{inact}$ values were calculated from the k_{inact} data in Tables 5.2 and 5.4 (which were calculated from Figures 5.5 and 5.8). The first two data columns show the difference in activation energies for the thermal inactivation process of *Ta*CS 5sm and *Ss*CS wt compared to *Ta*CS wt in the absence and presence of 18mM OAA and 0.2mM CoA. The positive values indicates an increase in thermal stability compared to *Ta*CS wt The final data column shows the difference in activation energies of the enzymes incubated in the presence of substrates compared to the enzymes incubated in the absence of substrates. A positive value indicates an increase in thermal stability in the presence of substrates.

5.4 Temperature optima of *Ta*CSs.

The temperature dependence of enzyme activity of *Ta*CS wt and the *Ta*CS ionic network mutants were determined by carrying out 2–minute enzyme assays at different temperatures between 40-85°C. Each measured rate of reaction was corrected for the degradation of DTNB, which increases with elevation in temperature resulting in an over-estimation of catalytic activity (~10% of actual enzymic rate at higher temperatures). Unfortunately, after prolonged storage at 4°C, *Ta*CS 6/3dm and *Ta*CS 3/5/6/tm had lost all activity prior to this experiment; therefore no temperature optima data are available for these mutants. This loss of activity on storage was not entirely unexpected, as these two *Ta*CS mutants were shown to be much less thermostable than the *Ta*CS wt (Section 5.3).

Figures 5.9-5.12 show the temperature dependence of the *Ta*CS ionic network mutants compared to both the *Ta*CS wt and the *Ss*CS wt. To aid identifying any changes in the temperature dependence data caused by any specific mutation, these data are displayed corresponding to the mutations they contain i.e. Figure 5.9 shows all mutants containing the M100R (RAM3) mutation and Figure 5.10 shows all mutants containing the T111E (RAM4) mutation etc.

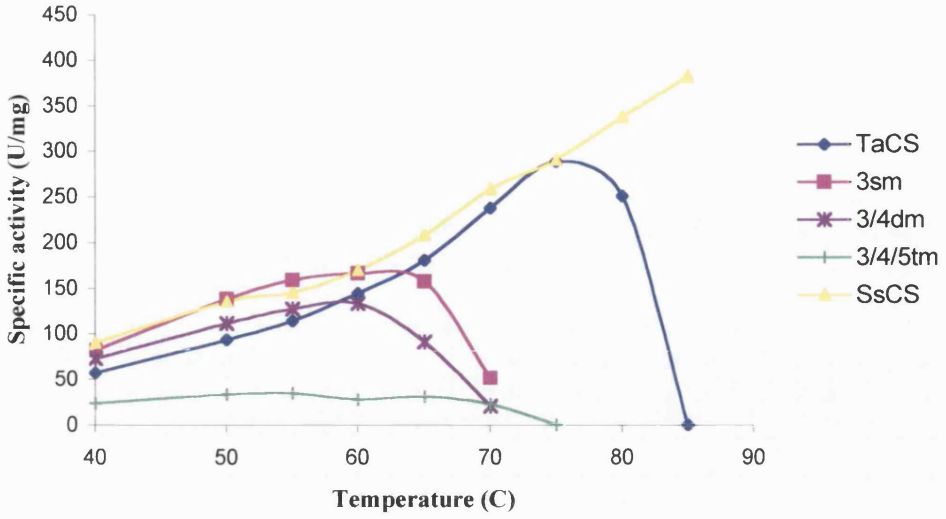


Figure 5.9 **Temperature dependence of the catalytic activity of *Ta*CS wt, *Ss*CS wt and all *Ta*CS ionic network mutants containing the M100R (RAM3) mutation.**
Enzyme assays were carried out at pH 8.0, at varying temperatures between 40-85°C. 0.14mM AcCoA and 1.25mM OAA were used in each assay, and initial rates were measured over 2 min.

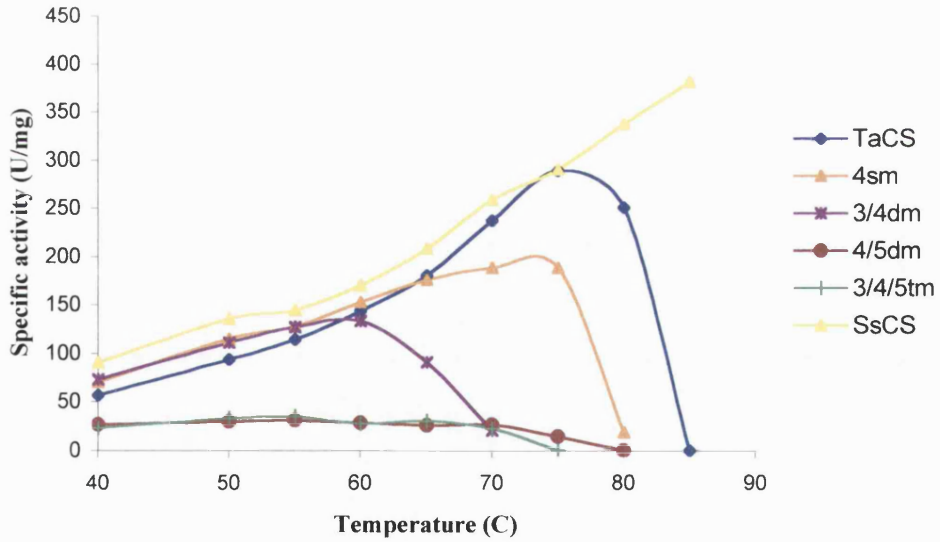


Figure 5.10 **Temperature dependence of the catalytic activity of *Ta*CS wt, *Ss*CS wt and all *Ta*CS ionic network mutants containing the T111E (RAM4) mutation.**
Enzyme assays were carried out at pH 8.0, at varying temperatures between 40-85°C. 0.14mM AcCoA and 1.25mM OAA were used in each assay, and initial rates were measured over 2 min.

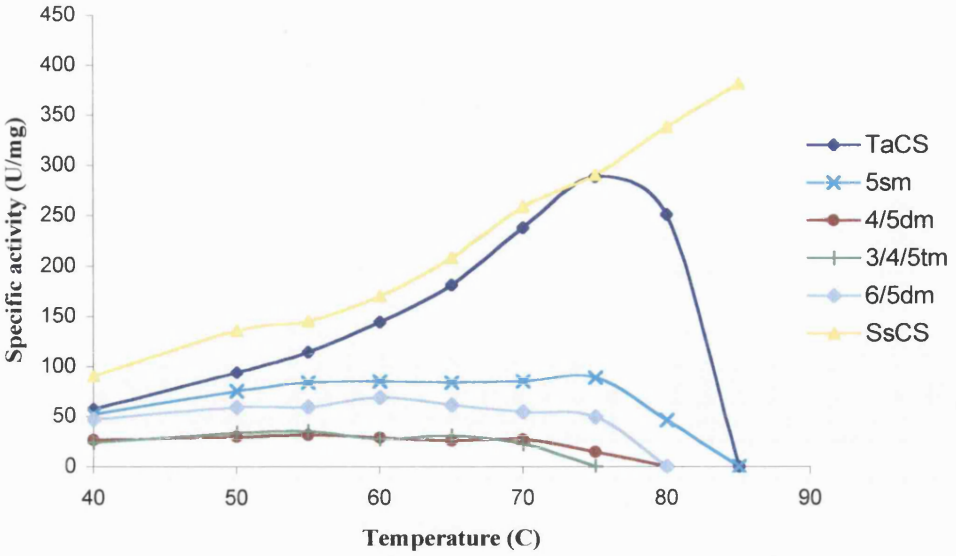


Figure 5.11 **Temperature dependence of the catalytic activity of *Ta*CS wt, *Ss*CS wt and all *Ta*CS ionic network mutants containing the D94H (RAM5) mutation.**
Enzyme assays were carried out at pH 8.0, at varying temperatures between 40-85°C. 0.14mM AcCoA and 1.25mM OAA were used in each assay, and initial rates were measured over 2 min.

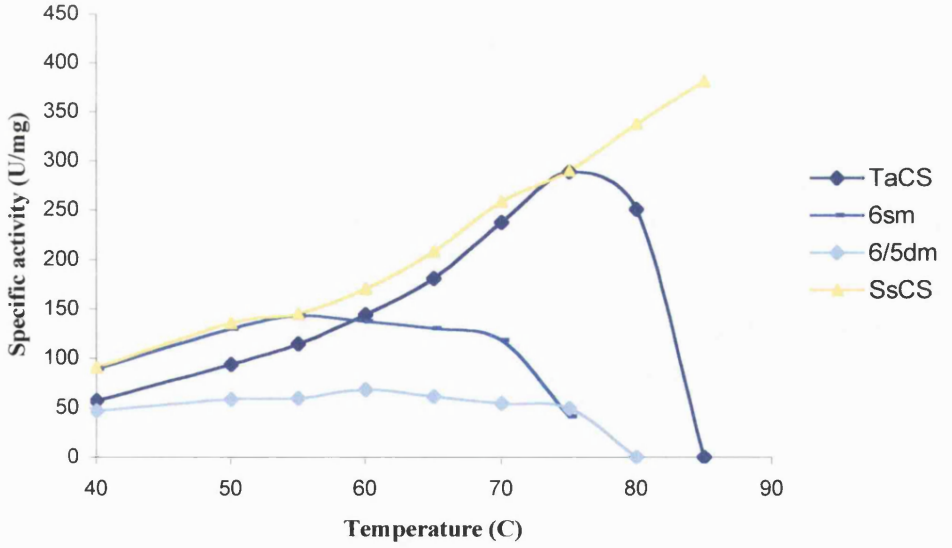


Figure 5.12 **Temperature dependence of the catalytic activity of *Ta*CS wt, *Ss*CS wt and all *Ta*CS ionic network mutants containing the G/D/I/P(RAM6) loop mutation.**
Enzyme assays were carried out at pH 8.0, at varying temperatures between 40-85°C. 0.14mM AcCoA and 1.25mM OAA were used in each assay, and initial rates were measured over 2 min.

These temperature dependence data (Figures 5.9-5.12) show that no *Ta*CS ionic network mutant has an increased temperature optimum (T_{opt}) compared to that of the *Ta*CS wt, (~75°C). As shown in Figure 5.11, all of the *Ta*CS mutants that contain the D94H (RAM5) mutation have an unusual flattened temperature dependence curve, which makes it difficult to estimate the T_{opt} value for these enzymes. It can be noted that, although there is no obvious T_{opt} value, all these mutants have a consistent specific activity over a range of temperatures (50-75°C), and only show a decrease in specific activity after 75°C, which is the temperature optimum of the *Ta*CS wt. These *Ta*CS mutants containing the D94H (RAM5) mutation also show a great reduction in the specific activity of the enzyme compared to the *Ta*CS wt. Figures 5.9, 5.10 and 5.12 show that only *Ta*CS 3sm, *Ta*CS 3/4dm, *Ta*CS 4sm and *Ta*CS 6sm have temperature dependence curves from which a T_{opt} value can be estimated.

Figures 5.13 and 5.14 show the temperature dependence data in the form of Arrhenius plots. These data are plotted according to the rearranged Arrhenius equation as stated in Section 3.3.4. Figure 5.13 illustrates the flattened nature of the temperature dependence curves for all of the *Ta*CS mutants containing the D94H (RAM5) mutation. It was difficult to obtain any activation energy (ΔG^*_{cat}) data from the slope of these plots, as the lines were too flat. Figure 5.14 shows the Arrhenius plot of the temperature dependence data excluding the *Ta*CS mutants with the D94H mutation. ΔG^*_{cat} values for these *Ta*CS ionic network mutants were determined from the linear portions of the plots in Figure 5.14, and are displayed in Table 5.6 along with the T_{opt} data for these mutants.

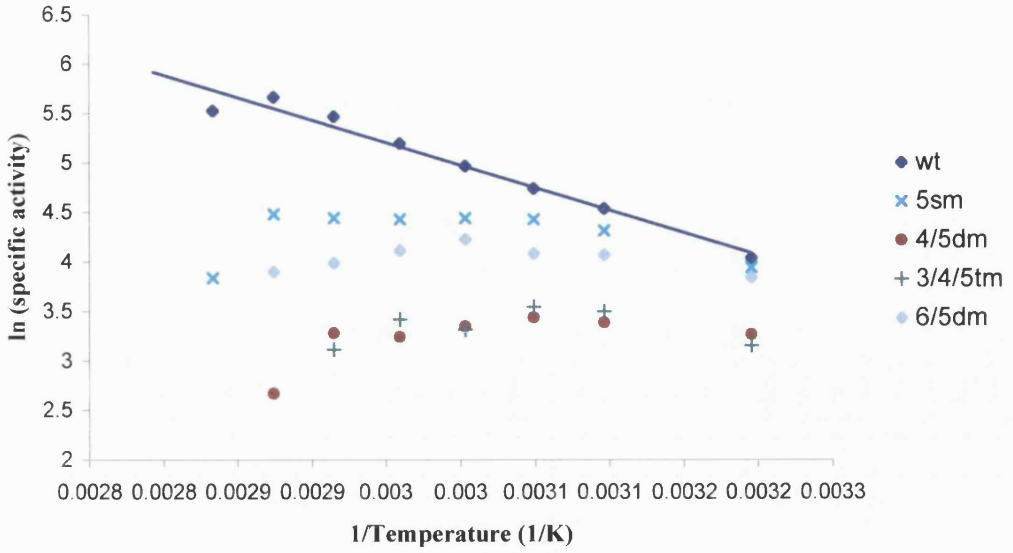


Figure 5.13 **Arrhenius plot of the catalytic activities of *Ta*CS wt, *Ss*CS wt and the *Ta*CS ionic network mutants containing the D94H (RAM5) mutation.**
Data shown are calculated from the temperature dependence data used for Figures 5.9-5.12. The line is fitted to the *Ta*CS wt data for illustrative purposes.

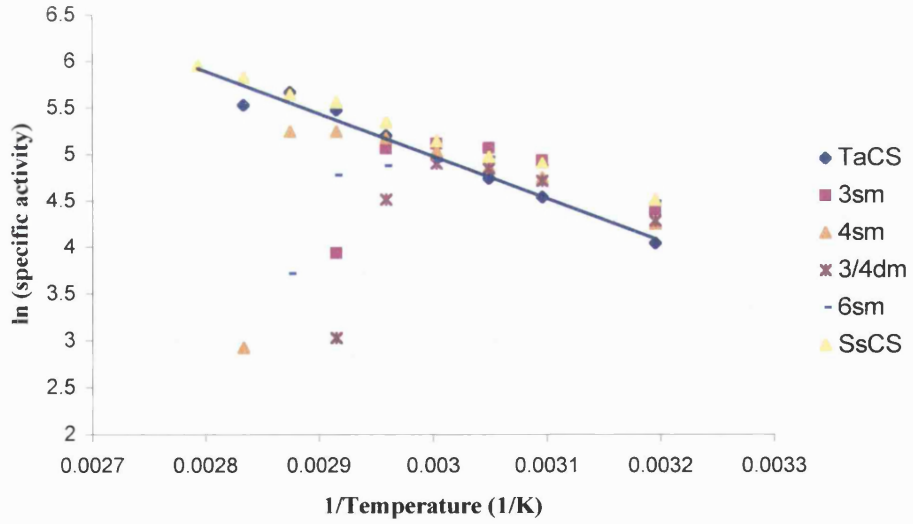


Figure 5.14 **Arrhenius plot of the catalytic activities of *Ta*CS wt, *Ss*CS wt and all the *Ta*CS ionic network mutants excluding any containing the D94H (RAM5) mutation.**
Data shown is calculated from the temperature dependence data used for Figures 5.9, 5.10 and 5.12. Line is fitted to the *Ta*CS wt data for illustrative purposes.

Citrate synthase	Temperature optimum, T_{opt} (°C)	ΔG^{*}_{cat} (kJ/mol)
*Ta*CS wt	75	42 (±1)
*Ta*CS 3sm	60	39 (±5)
*Ta*CS 4sm	70-75	29 (±2)
*Ta*CS 3/4dm	60	32 (±3)
*Ta*CS 6sm	55-60	28 (±4)
*Ss*CS wt	>85	30 (±1)

Table 5.6 **Temperature optima (Topt) and activation energies of the catalytic reaction (ΔG^{*}_{cat}) of *Ta*CS wt, *Ss*CS wt and all *Ta*CS ionic network mutants that do not contain the D94H (RAM5) mutation.**
The T_{opt} data are calculated from Figures 5.9, 5.10 and 5.12. The ΔG^{*}_{cat} data are calculated from the linear portion of the data in Figure 5.14.

Table 5.6 shows that the ΔG^{*}_{cat} values of the *Ta*CS ionic network mutants are decreased compared to the *Ta*CS wt, although there is less than a 1.5-2-fold decrease at the most. This suggests that the substitution of amino acids at the subunit interface of the *Ta*CS wt affects the activity and thermoactivity of this enzyme, which is consistent with the change in specific activity shown in Figures 5.9-5.12.

5.5 *Cis-trans* conformation of proline residues in the 4-membered loop region of *Ta*CS wt mutants

The crystal structure of *Pf*CS wt indicates that the proline residue in the loop region: Gly112, Asp113, Ile114 and Pro115 (G/D/I/P), is in the *cis* configuration (Figure 5.15(a)). This loop contains the aspartate residue necessary for the majority of the ionic interactions making up the ionic network present at the subunit interface (Chapter 3). This suggests that the *cis* configuration of this proline residue is important as it holds the loop near the subunit interface, hence, allowing the Asp113 residue to interact with other charged residues to form ionic bonds. This G/D/I/P loop region has been inserted into the equivalent position in the *Ta*CS wt (replacing Phe113 and Lys114).

Thermal characterization of the *Ta*CS ionic network mutants containing the G/D/I/P loop mutation (RAM6) shows a reduction in thermostability in each case, compared to the *Ta*CS wt (Section 5.3). This loop insertion mutation (RAM6) is an alternate mutation to RAM4, which is the substitution of Thr111 with a glutamate residue (T111E), see Chapter 4. Thermal inactivation data also shows that *Ta*CS mutants containing the RAM4 mutation show a decrease in thermostability compared the *Ta*CS wt. This may suggest that the introduction of these mutations is not resulting in the formation of the expected ionic interactions. In the case of the RAM4 (T111E) mutation, this may indicate that the glutamate residue may still not provide the sufficient length for the formation of ionic interactions in this region (Section 4.1). The data may also suggest that the inserted loop (RAM6) is not in the correct conformation to place the Asp113 in the correct position for interaction with the ionic network.

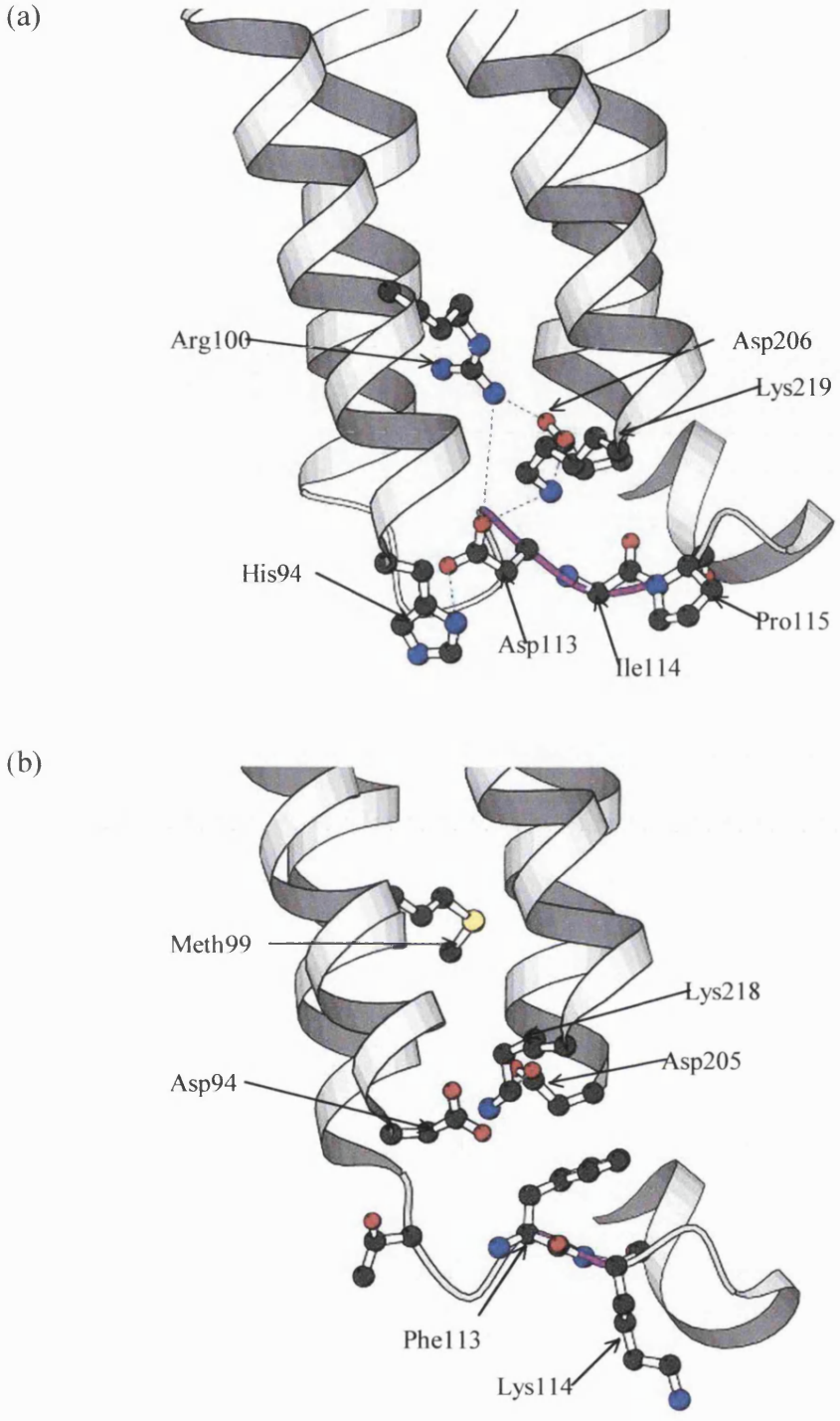


Figure 5.15 **Comparison of loop region in *Pf*CS and *Ta*CS**
Figure (a) shows the four amino acid loop region (Gly112, Asp113, Ile114, Pro115) in magenta. Figure (b) shows the equivalent residues in the *Ta*CS structure

Modeling of the *Ta*CS ionic network mutants containing the G/D/I/P loop insertion has suggested that the proline residue in this region may be in the *trans* conformation, as this is the most common conformation. This would result in the entire loop region being moved in the opposite direction to the intersubunit ionic network; hence prevent the formation of any ionic bonds. Study of the crystal structure of the *Ta*CS wt also suggests that there may be sufficient space in the protein for this proline residue to adopt a *cis* configuration. As there is a small, flexible glycine residue on the opposite side of the loop region (Gly112), it may be possible to 'flip' the loop into the correct position. This would require the presence of an enzyme that can catalyze the *cis/trans* isomerisation of proline residues.

5.6 Peptidyl proline *cis-trans* isomerase (PPIase)

5.6.1 Introduction to PPIase.

In proteins, the partial double bond character of the peptide bond results in two conformations depending on the value of the dihedral angle: *cis* and *trans* with ω=o and ω=180°, respectively [Stewart *et al.*, 1990], see Figure 5.16(a). The isomer with the two C^{α} atoms *trans* to each other is favoured due to the lesser steric conflict involving the substituents at these positions. A difference in energy of approximately 20kJ/mol (5kcal/mol) exists between the *trans* and *cis* isomers, and therefore, *cis-trans* isomerisation of a peptide bond is a slow process at room temperature and has shown to play an important role in protein folding (recently reviewed by Pal & Chakrabati., 1999 and Eyles & Gierasch., 2000). However, *cis* and *trans* conformations of X-proline

bonds (where X represents any amino acid) have comparable energies (80-100kJ/mol), the *trans* form being favoured by ~4kJ/mol (1kcal/mol).

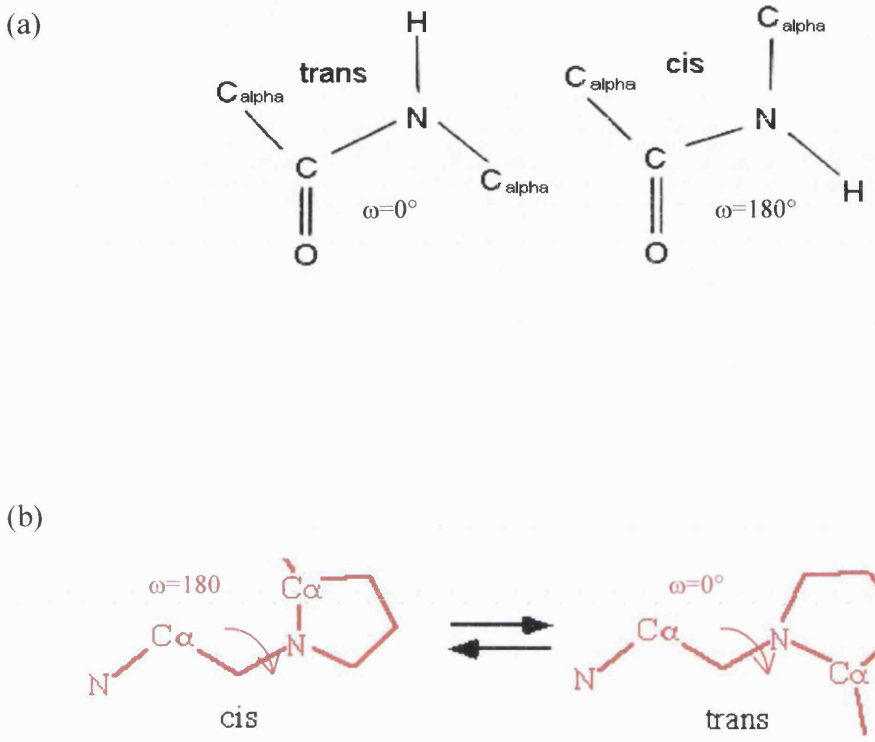


Figure 5.16 ***cis/trans* isomerisation of a peptidyl-prolyl bond.**
Figure (a) shows the conformations of *trans* and *cis* isomers of a general peptide bond. Figure (b) shows both the *cis* and *trans* conformers of a peptidyl-prolyl bond.

Consistent with these energetic data, a survey of protein structures by Stewart *et al.* (1990) found only 0.05% of all X-Xnp (where Xnp is also any other amino acid), but 6.5% of all X-Pro peptide bonds, occur in the *cis* conformation in the folded protein

(Figure 5.16(b)). *Cis* peptidyl-prolyl bonds are found primarily in bends and turns, suggesting a specific structural role for this type of bonding [Stewart *et al.,* 1990].

It has been suggested that, as proline residues reduce the conformational degrees of freedom in the main polypeptide chain, they may be involved in increasing the thermostability of a protein by increasing the rigidity of the polypeptide chain [recently reviewed by Watanabe *et al.*, 1997]. Watanabe *et al.* (1997) showed that introducing proline substitutions in the homologous oligo-1,6-glucosidases from various *Bacillus* strains growing within different temperature ranges, identified two specific proline sites critical to protein thermostabilization. Bogin *et al.* (1998) compared the crystal structures of the closely related mesophilic *Clostridium beijerinckii* alcohol dehydrogenase (CBADH) and the hyperthermophilic *Thermoanaerobacter brockii* alcohol dehydrogenase (TBADH), and suggested that the extra proline residues located in strategically important positions in TBADH might contribute to the thermal stability of this enzyme. Site-directed mutagenesis was used to replace eight complementary residue positions in CBADH, one at a time, with proline residues. They showed that two of the eight proline residues in TBADH contribute to its thermal stability.

The *cis/trans* isomerisation of proline-peptide bonds is often the rate-limiting step in protein folding. Enzymes that catalyze this step are called peptidyl-prolyl *cis/trans* isomerases (PPIases). PPIases are ubiquitous in living organisms and three families, cyclophilin (CyP), FFK506 binding protein (FKBP) and parvulin (Pvn), have been studied in detail (recently reviewed by Gothel *et al.*, 1999 and Maruyama *et al.*, 2000). The cyclophilins form a family of cellular proteins that bind the drug cyclosporin A. For this reason they are also called immunophilins. *In vitro* cyclophilins accelerate slow

protein folding reactions by catalyzing the cis/trans isomerisation of peptide bonds preceding a prolyl residue.

5.6.2 PPIase experiments.

For these experiments calf thymus Cyclophilin (CyP) PPIase was used (Sigma, product number C7696). Due to time constraints on this project, it was impossible to design and carry out unfolding (in guanidine or urea) and refolding experiments for the *Ta*CS ionic mutants. Instead, it was decided to try to change the conformation of the proline (on the G/D/I/P inserted loop, RAM6) from *trans* to *cis* on the intact protein. This enzyme has previously been used in *in vivo* protein folding experiments, but there a no evidence that it has ever been used to catalyze the conversion of a *trans* X-Pro bond to a *cis* conformation in a folded protein.

These experiments consisted of incubating both *Ta*CS 6/5dm and *Ta*CS 3/5/6tm separately in buffer (pH 7.0) containing the cyclophilin enzyme. These *Ta*CS mutants were chosen because they contain both the G/D/I/P loop mutation (RAM6) and other amino acids that then give the potential of making one or more ionic interactions. On the other hand, the *Ta*CS 6sm containing just the loop mutation was not investigated as it does not contain the other residues required to form ionic interactions (see Figure 4.2). Similarly, *Ta*CS 6/3dm was not chosen either as it has been shown to be very unstable (Sections 5.3 and 5.4) with a greatly reduced thermal stability compared to the *Ta*CS wt. This may be due to the M100R (RAM3) mutation that has previously been shown to be a destabilizing mutation.

Published protocols for the use of PPIases have all involved the renaturing of denatured proteins or the effects on small peptides. Fischer (1985) showed that the optimum pH for PPIase activity when refolding urea-denatured ribonuclease A was pH 7.8. Schiene *et al.* (1998) used 0.1-5μM of various PPIases (including human cyclophilin18) on 400μM of a tetrapeptide substrate (a PPIase:substrate ratio of 1:4000- 1:80) to map the stereospecificity (pH 7.8). von Ahsen *et al.* (2000) used 1.5μM of various cyclophilins (from yeast, *Neurospora crassa* and *E.coli*) for the refolding of 15μM guanidine-denatured mouse dihydrofolate reductase, DHFR (pH 7.2), (a PPIase: substrate ratio of 1:10). It must be noted that these are high concentrations of cyclophilin used, as calf thymus cyclophilin has a molecular weight of 17kDa. A 1μM solution of calf thymus cyclophilin gives a protein concentration of 0.03mg/ml, which is the same concentration as the solutions of *Ta*CS 6/5dm and *Ta*CS 3/5/6tm used for the thermal inactivations (a 1:1 ratio), (Section 5.3). Many enzyme reactions i.e. restriction digests etc., use <10% concentration of enzyme to substrate. For this reason, *Ta*CS 6/5dm and *Ta*CS 3/5/6tm were incubated with 1μM of cyclophilin (1:1 ratio), but also this incubation was repeated with 1:10 and 1:50 cyclophilin:substrate.

Most of the protein refolding experiments using a PPIase in the literature carried out their experiments at room temperature (25°C), but as the cyclophilin used here is from calf thymus then incubation at 37°C was also used. All incubations were left overnight. The thermal inactivations of the incubated samples were carried out as before (Section 5.3). The initial specific activities of the citrate synthases were not affected by the incubation with cyclophilin.

Figure 5.17 shows the thermal inactivation at 80°C of *Ta*CS 6/5dm and the *Ta*CS wt, in the absence and presence of 1μM cyclophilin incubated at either 25°C or 37°C overnight. The thermal inactivation was done at 80°C as the *Ta*CS 6/5dm still has activity remaining (12%) after a 25 minute incubation at this temperature (see Figure 5.4, Section 5.3). This figure shows that the incubation of the *Ta*CS wt and *Ta*CS 6/5dm with cyclophilin has no effect on the thermostability of the enzyme, whether incubated at 25°C or 37°C.

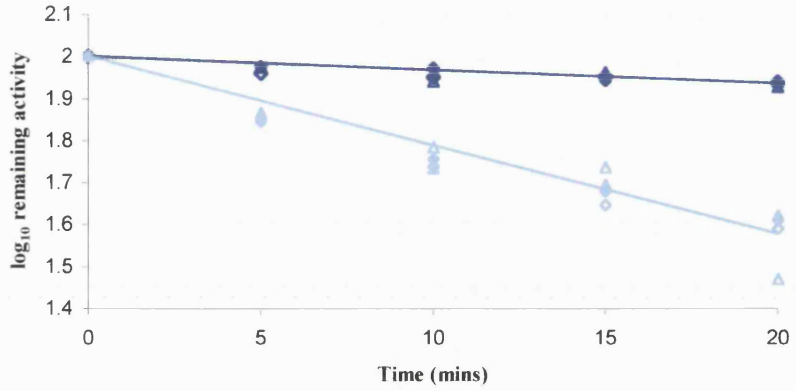


Figure 5.17 **Thermal inactivation of *Ta*CS wt and *Ta*CS 6/5dm at 80°C after incubation in the presence and absence of cyclophilin, at 25 or 37°C.**
0.03μg/ml solutions of *Ta*CS wt and *Ta*CS 6/5dm were incubated at pH 7.0, with and without 1μM cyclophilin (CyP), 1:1 ratio of cyclophilin:substrate, overnight at 25 or 37°C. After the overnight incubation, the enzymes were incubated at 80°C, and samples were removed at 5-minute intervals and cooled in an ice-water bath. These samples were assayed as standard at 55°C. Each *Ta*CS wt sample: incubated at 37°C w/o CyP (◆), incubated at 37°C with CyP (◇), incubated at 25°C w/o CyP (▲) and incubated at 25°C with CyP (△), showed a similar thermal inactivation. Each *Ta*CS 6/5dm sample: incubated at 37°C w/o CyP (◆), incubated at 37°C with CyP (◇), incubated at 25°C w/o CyP (▲) and incubated at 25°C with CyP (△), showed the same degree of thermal inactivation relative to the *Ta*CS wt. The lines are best fits to the data for the incubation at 37°C in the absence of 1μM CyP for each enzyme.

Figure 5.18 shows the thermal inactivation at 80°C of the *Ta*CS 6/5dm and the *Ta*CS wt incubated, overnight at 37°C, with either 1:10 or 1:50 cyclophilin:substrate. Figure 5.19 shows the thermal inactivation at 70°C of the *Ta*CS 3/5/6tm and *Ta*CS wt incubated, overnight at 37°C, with 1:10 and 1:50 cyclophilin:substrate. Again, 70°C was used as the incubation temperature as Figure 5.2 (Section 5.3) shows that *Ta*CS

3/5/6tm still has activity remaining (20%) after a 10 minute incubation at this temperature.

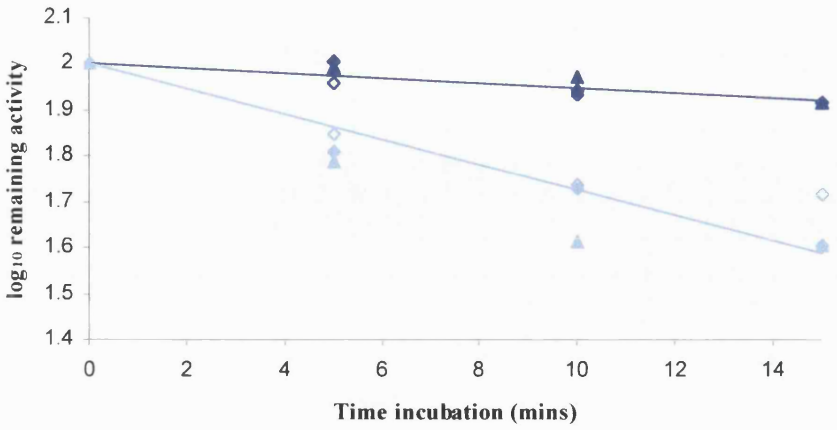


Figure 5.18 **Thermal inactivation of *Ta*CS wt and *Ta*CS 6/5dm at 80°C after incubation in the presence and absence of 1:10 or 1:50 cyclophilin:substrate concentration, at 37°C.** 0.03μg/ml solutions of *Ta*CS wt and *Ta*CS 6/5dm were incubated at pH 7.0, with and without 1:10 or 1:50 cyclophilin(CyP):substrate, overnight at 37°C. After the overnight incubation, the enzymes were incubated at 80°C, and samples were removed at 5-minute intervals and cooled in an ice-water bath. These samples were assayed as standard at 55°C. Each *Ta*CS wt sample: incubated w/o CyP (◆), incubated at 1:10 CyP (◇) and incubated at 1:50 CyP (▲), showed a similar thermal inactivation. Each *Ta*CS 6/5dm sample: incubated w/o CyP (◆), incubated at 1:10 CyP (◇) and incubated at 1:50 CyP (▲), showed the same degree of thermal inactivation relative to the *Ta*CS wt. The lines are best fits to the data for the incubation in the absence of cyclophilin for each enzyme.

Figures 5.17-5.19 all show that incubation of the *Ta*CS ionic network mutants containing the G/D/I/P inserted loop, with cyclophilin at varying concentrations, does not effect the thermostability of the mutants. This is unsurprising as the proteins are already folded which would make it very difficult to change the conformation of one amino acid within the protein structure. However, these data do not show whether or not the proline bond has undergone a change in *cis/trans* configuration, just that there is no observed change in the thermostability of these enzymes.

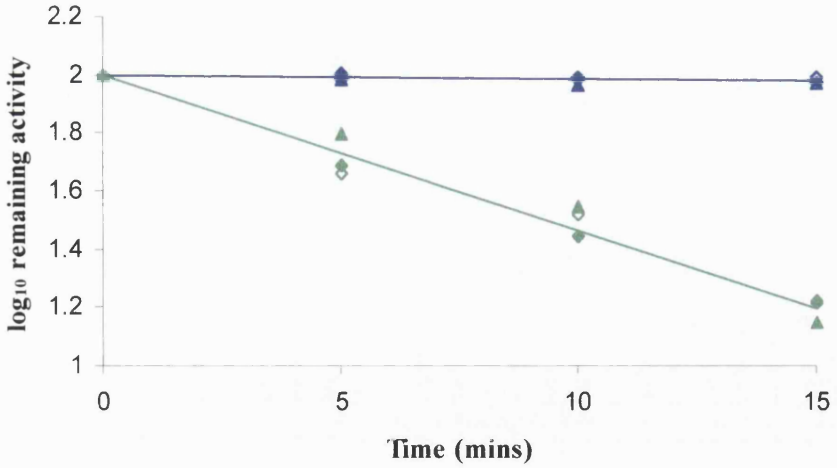


Figure 5.19 **Thermal inactivation of *Ta*CS wt and *Ta*CS 3/5/6tm at 70°C after incubation in the presence and absence of 1:10 or 1:50 cyclophilin:substrate concentration, at 37°C.** 0.03µl/ml solutions of *Ta*CS wt and *Ta*CS 3/5/6tm were incubated at pH 7.0, with and without 1:10 or 1:50 cyclophilin(CyP):substrate, overnight at 37°C. After the overnight incubation, the enzymes were incubated at 70°C, and samples were removed at 5-minute intervals and cooled in an ice-water bath. These samples were assayed as standard at 55°C. Each *Ta*CS wt sample: incubated w/o CyP (◆), incubated at 1:10 CyP (◇) and incubated at 1:50 CyP (▲), showed a similar thermal inactivation. Each *Ta*CS 3/5/6tm sample: incubated w/o CyP (◆), incubated at 1:10 CyP (◇) and incubated at 1:50 CyP (▲), showed the same degree of thermal inactivation relative to the *Ta*CS wt. The lines are best fits to the data for the incubation in the absence of cyclophilin for each enzyme.

It may be interesting to vary the conditions of the incubation with cyclophilin; an increased time of incubation or variation in temperature could be employed. However, these results suggest that the Pro115 residue may only be converted from *trans* to *cis* during an unfolding and refolding experiment in the presence of a PPIase. Unfortunately, time did not permit these experiments to be carried out, although such investigations are on-going in the research group.

5.7 Discussion

Kinetic characterization of the *Ta*CS ionic network mutants shows that, at 55°C, there is an increase in the K_m for AcCoA for all of the mutant enzymes compared to *Ta*CS wt, some increases being as large as 10-fold (*Ta*CS 4/5dm) (Table 5.1). There is also an

overall increase in the K_m values for OAA (apart for *Ta*CS 3/4/5tm and *Ta*CS 6sm), although no more than 2.5-fold compared to the wild-type citrate synthase. Most of the *Ta*CS mutants have a V_{max} value comparable to that of the *Ta*CS wt (70-80 μmol/min/mg), although both of the mutants containing the full ionic networks (*Ta*CS 3/4/5tm and *Ta*CS 3/5/6tm) have only 52 and 66% the catalytic activity of *Ta*CS wt respectively. Changes in enzyme activity are not unexpected as mutations at the subunit interface may have an effect on the structure of the citrate synthase active site, which lies between the two monomers. The data suggest that, in some cases, the active site might be affected by mutations, but no intersubunit mutation has destroyed the catalytic activity completely. This is similar to the situation with *Pf*CS where mutations to disrupt the inter-subunit ionic network had little effect on catalytic activity [Arnott *et al.*, 2000], as discussed in Chapter 3.

Interestingly, some of ionic network mutants show a large increase in V_{max} at 55°C, most noticeably for *Ta*CS 3sm and *Ta*CS 3/4dm at pH 8.0 (Table 5.1). This increase in enzyme activity at lower temperatures has also been noted in the thermal inactivation and temperature dependence data obtained (Table 5.2, Figure 5.9). These data suggest that some mutations at the intersubunit region of *Ta*CS wt cause an increase in catalytic activity at lower temperatures. This may result from a loss of rigidity of the citrate synthase enzyme i.e. an increase in the degree of flexibility. It has been shown that an increase in thermostability may coincide with an increase in rigidity; hence a loss in rigidity can result in a loss of thermal stability but an increase in activity at lower temperatures.

For ease of discussion, thermal inactivation and temperature dependence data are discussed with respect to each of the four mutations involved in constructing the five-membered ionic network into *Ta*CS wt, M100R (RAM3), T111E (RAM4), D94H (RAM5) and the G/D/I/P loop insert (RAM6). A reproduction of Figure 4.2 from Chapter 4 is shown here as Figure 5.20 for reference purposes.

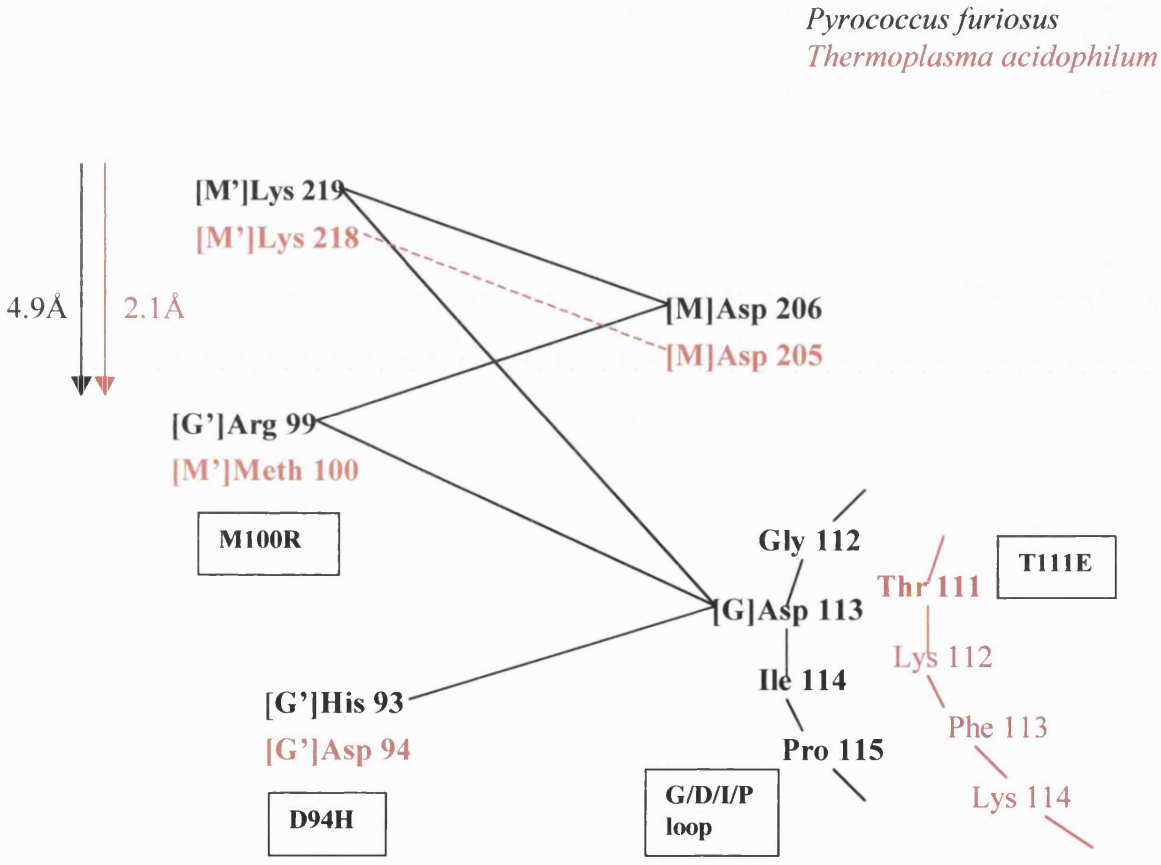


Figure 5.20 **Schematic comparison of ionic network regions of *Ta*CS wt and *Pf*CS wt**
T.acidophilum residues are shown in red and *P.furiosus* residues are shown in black. Each change of residue between the two citrate synthases is named corresponding to Table 4.2. This figure is a reproduction of Figure 4.2

5.7.1 Methionine100 –arginine, M100R (RAM3).

The first mutation introduced into the *Ta*CS wt subunit interface was the substitution of Meth100 with an arginine residue. This mirrors the Arg99 residue in the ionic network of the *Pf*CS wt, which forms ionic bonds with Asp206 and Asp113 on the opposite subunit (Figures 5.20 and 5.21).

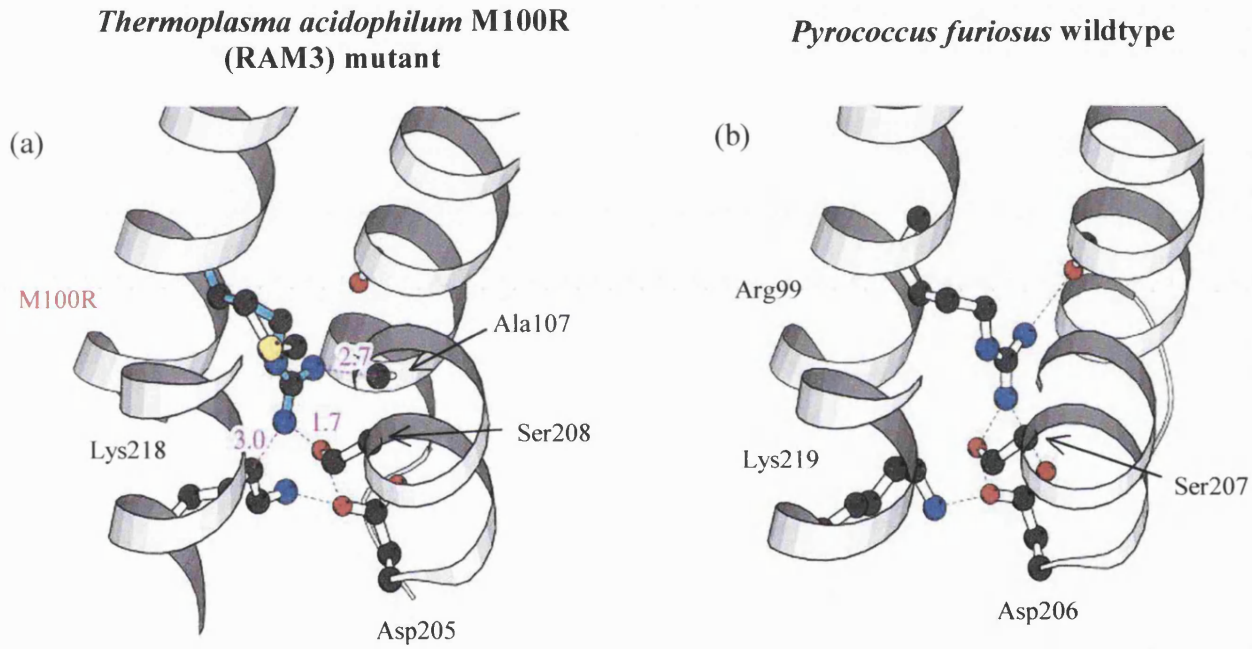


Figure 5.21 **Predicted structure of the subunit interface of *Ta*CS 3sm compared to the crystal structure *Pf*CS wt.**
For the M100R (RAM3) substitution (Figure (a)), the main chain of the arginine residue is aligned with the first two side chain atoms of the Meth100 residue, thereby keeping its conformation the same as in *P.furiosus* citrate synthase (Figure (b)). Normal hydrogen bonds are shown in green and the clashes are shown in magenta.

Thermal inactivation and temperature optima data show that there is a distinct decrease in the thermostability of all the *Ta*CS mutants containing the M100R mutation. Figure 5.21(a) shows the M100R mutation modeled into the *Ta*CS wt, compared to the ionic network of *Pf*CS wt (Figure 5.21(b)). In Figure 5.21(a) the arginine residue is introduced into the *Ta*CS by aligning its main chain and the first two side chain atoms with the Meth100, thereby keeping its conformation the same as in *Pf*CS wt. This model shows that the M100R mutation results in 3 clashes with neighbouring residues. Firstly, the introduced arginine residue might be too close to the Ser208 residue, ~1.7Å, which is almost the same length as a carbon-carbon bond. Also, the substituted residue may clash, both sterically and with opposing charge, with Lys218 and Ala107. If this model were correct then these clashes would cause the helices of the subunit interface to move away from each other, which may explain the decrease in thermostability of enzymes containing this mutation. This may also introduce flexibility in the interface region of the citrate synthase enzyme, which may explain the increase in catalytic activity of these mutants at lower temperatures.

5.7.2 Threonine111-glutamate, T111E (RAM4)

Chapter 3 discusses the substitution of Asp113 in the ionic network at the subunit interface of *Pf*CS wt. The equivalent residue in the *Ta*CS wt is Thr111. To overcome this problem it was decided to substitute the Thr111 with a glutamate residue as the glutamate is of equivalent charge to the aspartate but is a carbon bond length longer.

Figure 5.22 (a) shows the predicted position of the T111E mutation in the *Ta*CS wt subunit interface. This figure illustrates the glutamate residue in its most favourable

rotamer, and shows that it has the potential to form ionic bonds including both its oxygen atoms. Figure 5.22(b) shows that at the subunit interface of *Pf*CS wt the Asp113 forms an ionic bond, 3.05Å, with Lys219. The interactions of the glutamate residue in the TaCS T111E mutant with Lys218 are much longer, about 6Å, and hence potentially much weaker.

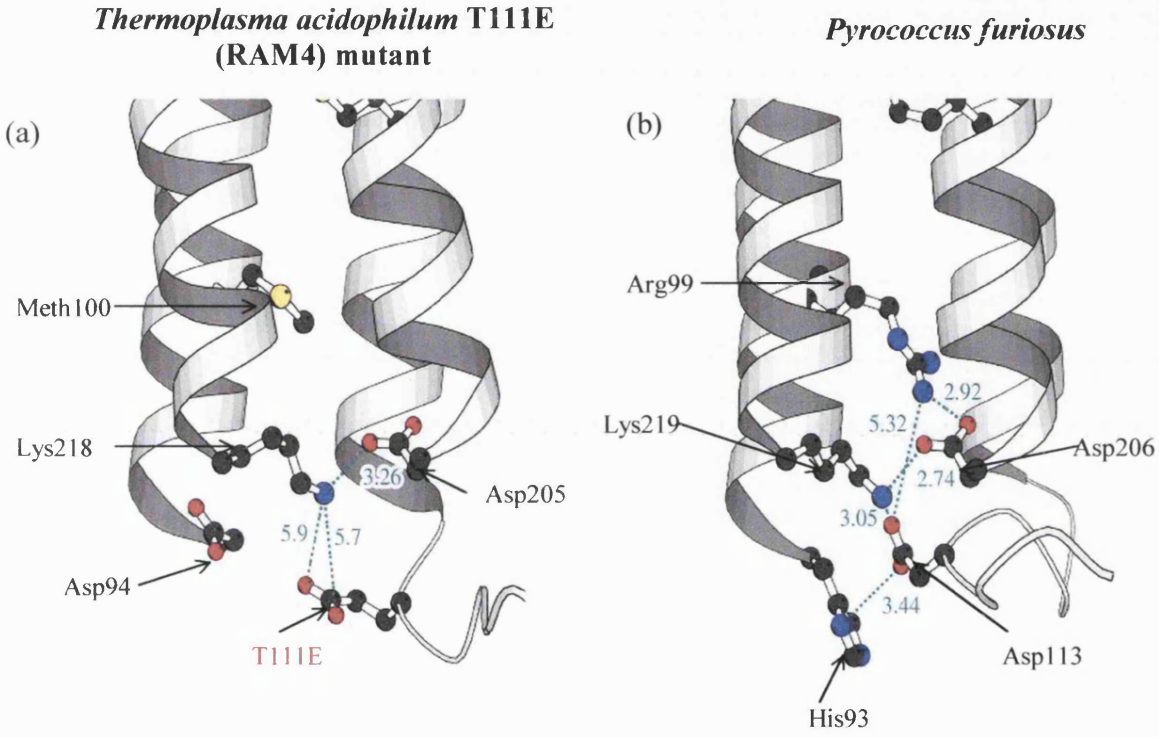


Figure 5.22 **Predicted structure of subunit interface of the TaCS 4sm compared to the crystal structure of *Pf*CS wt.**
Figure (a) shows the T111E substitution positioned taking the most favourable rotamer of glutamate to possibly make 2 long ionic bonds with Lys218.

Hence, Figure 5.22(a) shows that the T111E mutation would have little effect in increasing the thermostability of the *Ta*CS enzyme. This may explain the fact that the *Ta*CS 4sm mutant shows little change in kinetic parameters or T_{opt}, so no change in the structure of the active site, but the thermal stability is affected suggesting that this mutation has introduced a small structural change to the enzyme. Hence, any change in

the thermostability of the *Ta*CS mutants containing the T111E mutation, in combination with other mutations, may be a result of the effect of the other mutations.

5.7.3 Aspartate94-histidine, D94H (RAM5).

This mutation is a direct substitution of a histidine residue for Asp94 in the *Ta*CS wt to mirror the His93 in the intersubunit ionic network of *Pf*CS wt. Kinetic data indicate that, apart from the full network mutants (*Ta*CS 3/4/5tm and *Ta*CS 3/5/6tm) that have a reduced V_{max}, this D94H (RAM5) mutation does not have a great effect on the V_{max} values (Table 5.1). This mutation does appear to increase the Km for both AcCoA and OAA.

Thermal inactivation data, in the absence of substrates (Figures 5.1-5.5), show that apart from the full network mutants, *Ta*CS 3/4/5tm and *Ta*CS 3/5/6tm, all of the *Ta*CS mutants containing the D94H (RAM5) mutation retain much of their thermostability. The double mutants, *Ta*CS 4/5dm and *Ta*CS 6/5dm, have $t_{½}$ values of 2.8 and 8.3 minutes at 85°C and 80°C respectively (Table 5.2). Although these *Ta*CS ionic network mutants are not as thermostable as the *Ta*CS wt, they are more thermostable than the other ionic network mutants constructed. Temperature dependence data (Figure 5.12) show that all of the mutants containing the D94H mutation have an unusual, flat temperature dependence curve with no clear temperature optima (T_{opt}). These mutants do not appear to lose activity until 70-75°C, similar to the T_{opt} of the wild-type, but do not have an increase in activity with increasing temperatures up to this temperature.

The most interesting mutation constructed appears to be the *Ta*CS ionic network mutant containing the D94H mutation alone, *Ta*CS 5sm. Thermal inactivation at 85°C, in the absence of substrate (Figure 5.5), shows that this mutant is more thermostable than the *Ta*CS wt i.e. *Ta*CS 5sm has a 2-fold decrease in k_{inact} compared to the *Ta*CS wt (Table 5.2). This increase in thermostability becomes more interesting when the thermal inactivation is carried out in the presence of CoA and OAA (Figure 5.8). This figure shows that the *Ta*CS 5sm mutant is more thermostable than the *Ta*CS wt, but it also shows that this mutant has a thermal stability, at 85°C, comparable to the *Ss*CS wt. Interestingly, Table 5.5 indicates that both the *Ta*CS wt and the *Ta*CS 5sm mutant have an increase in thermostability when incubated in the presence of substrates, 2.4 and 3.2 KJ/mol respectively.

Figure 5.23 illustrates a speculative arrangement of the added histidine residue in place of the Asp94 in the *Ta*CS wt. The histidine was overlaid onto the first two carbons of the Asp94, and then orientated in the equivalent position as in *Pf*CS wt. The magenta interaction shows that the histidine is closer to Lys218, 4.6Å, than the Asp94-Lys218 interaction in the *Ta*CS wt. This may result in the clashing charges pushing the histidine residue away from the positively charged lysine residue. This movement in residues may cause a weak ionic interaction with the negatively charged Glu92 (Figure 5.23 (a)); this interaction cannot exist in the *Ta*CS wt, as the Asp94 is a negatively charged residue (Figure 5.23(b)). The ionic interaction formed between the histidine and Glu92 may pull the glutamate residue closer to the histidine residue, which would put the Glu92 in a position to form an ionic bond with Lys112.

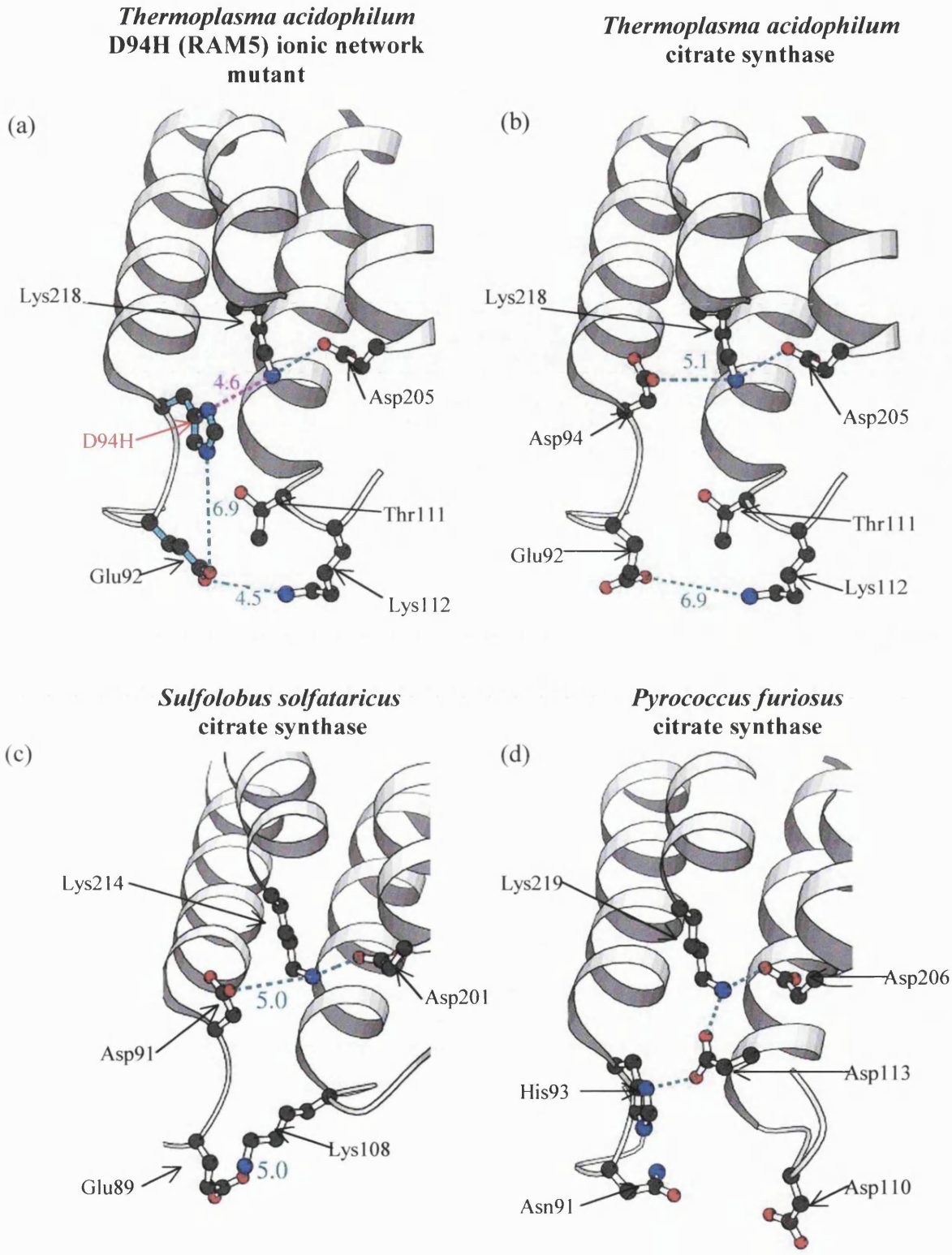


Figure 5.23 **Comparison of the predicted structure of the *Ta*CS 5sm mutant to the *Ta*CS wt, *Ss*CS wt and *Pf*CS wt structures.**
Figure (a) shows the predicted structure of *Ta*CS 5sm The histidine residue is overlaid on the first two carbon atoms of the Asp94 that is shown in Figure (b). The magenta interaction shows how this mutation is closer to the Lys218 residue. Figures (b), (c) and (d) show the equivalent region in *T.acidophilum, S.Sulfolobus* and *P.furiosus* citrate synthases.

This bond is not present in the PfCS wt (5.23(d)). The Glu92-Lys112 bond that is formed by this mutation is equivalent to the second ionic bond that is present in the ionic network region of *Ss*CS wt, Glu98-Lys108, 5.1Å, (Figure 5.23(c)). This may explain why initial characterization of the *Ta*CS 5sm mutant shows it to be more thermostable than the *Ta*CS wt, and has equivalent thermal stability to the *Ss*CS wt in the presence of substrates

5.7.4 Four-residue loop insertion, G/D/I/P (RAM6)

As described in Chapter 4, Section 4.1, and Section 5.5 of this chapter, the Asp113 residue of the ionic network in the *Pf*CS wt is positioned on a four-residue loop at the subunit interface (Figure 5.24). This loop, consisting of Gly112, Asp113, Ile114 and Pro115 (G/D/I/P), is not present in the *Ta*CS wt. This results in the Thr111 residue in *Ta*CS wt, the equivalent residue to Asp113, being positioned too far away from the ionic network region to be substituted for an aspartate that will be able to interact with other members of the ionic network. The T111E (RAM4) mutation attempts to shorten this distance by substituting in a glutamate, which is longer than the aspartate residue, but this mutation did not work sufficiently well (Section 5.7.2). The alternative method was to insert the entire four-residue loop, G/D/I/P (RAM6), into the subunit interface of *Ta*CS wt.

Kinetic data (Table 5.1) indicate that the mutant containing the G/D/I/P loop (RAM6) alone shows very little change in the kinetic parameters compared to *Ta*CS wt. No temperature dependence data was obtained for these mutants, as all enzyme activity was

lost, in each case, after extended storage at 4°C. This also highlights the instability of these enzymes in general.

As discussed previously, it is possible that the T111E (RAM4) mutation failed, as it was unable to make an ionic interaction due to its distance from the ionic network region (Section 5.7.2). The insertion of the loop (RAM6) should be able to decrease the distance of the inserted Asp113 residue from the subunit interface. Section 5.5 of this chapter discusses the possibility that the proline residue in this inserted loop region may be in the incorrect conformation. In *Pf*CS wt, the Pro115 residue is in the *cis* conformation, which is the more unfavourable conformation (Section 5.6.1), and Figure 5.24(b) shows how the *cis* conformation of the proline residue is essential for positioning the Asp113 correctly. It is possible that the proline in the inserted loop in the *Ta*CS mutants is in the *trans* conformation, pushing the Asp113 residue away from the ionic network region.

Section 5.6 describes experiments carried out in an attempt to convert the possible *trans* proline residue into a *cis* conformation in the *Ta*CS ionic network mutants containing the G/D/I/P loop (RAM6) mutation using a peptidyl proline *cis-trans* isomerase (PPIase). Due to time restrictions it was impossible to unfold the enzymes then refold them in the presence of the PPIase, so the enzymes were just incubated with the PPIase overnight and then thermally inactivated at 75 and 80°C.

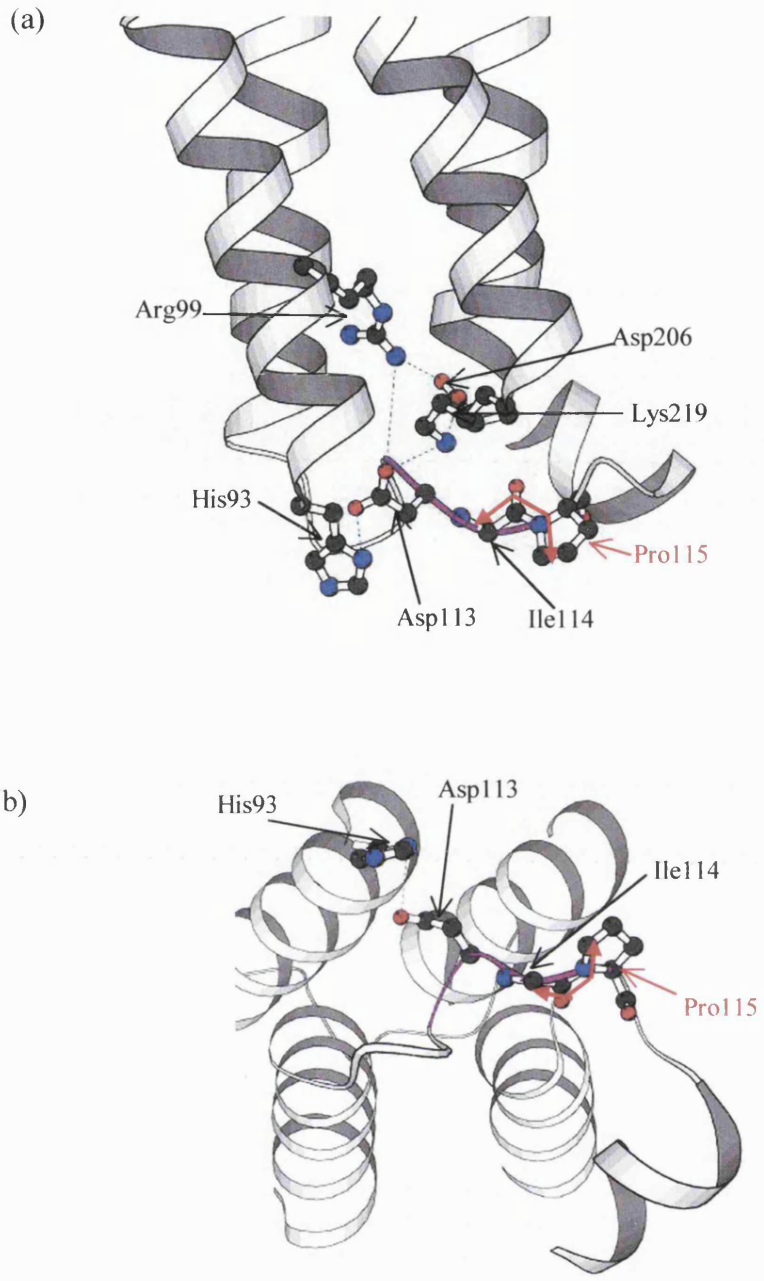


Figure 5.24 **The intersubunit loop region in the *Pf*CS wt.**
Figure (a) depicts the five-membered ionic network at the subunit interface of *Pf*CS wt, including the four-residue loop region that is responsible for positioning Asp113 so that ionic bonds may be formed. The *cis* conformation of the Pro115 residue is highlighted in red. Figure (b) shows the same structure from a position that shows how the *cis* prolyl bond is essential for the positioning of the Asp113 residue. Again, the *cis* prolyl bond is highlighted in red. The entire loop region is shown in magenta in both figures.

These experiments were unsuccessful in increasing the thermostability of *Ta*CS 6/5dm and *Ta*CS 3/5/6tm, although the results did not indicate whether or not the proline residue had changed conformation, the data only showed that there was no resulting increase in thermostability. It would be more beneficial to unfold the *Ta*CS ionic network mutants before refolding in the presence of the PPIase. It would also be interesting to obtain crystal structure data for these mutants to discover the actual orientation of the proline residue in the inserted loop region.

In conclusion, attempts to increase the thermostability of *Ta*CS wt by mirroring the five-membered ionic network present in the *Pf*CS wt, using site-directed mutagenesis, seem to have been unsuccessful. Thermal characterization shows that only one *Ta*CS ionic network mutant, *Ta*CS 5sm, has any increase in thermostability compared to the *Ta*CS wt. Interestingly, this *Ta*CS mutant contains only a single mutation that, as modeling studies suggest, may form an interaction that mirrors the second ionic bond found at the subunit interface of the more thermostable *Ss*CS wt.

Clearly, it is necessary to obtain crystallographic data for the *Ta*CS 5sm to determine the structural changes of the amino acid substitution for comparison to the modeling data. Further thermal characterization of this *Ta*CS mutant is also required to study the extent of the increase in the thermostability compared to *Ta*CS wt and *Ss*CS wt.

CHAPTER 6

Discussion

6.1 The role of the ionic network of *Pf*CS in thermostability

This site-directed mutagenesis study of the 5-membered ionic network present at the subunit interface of the *Pf*CS has shown that substitution of the Asp113 residue with a serine or alanine residue reduces the thermal stability of the enzyme (Chapter 3). This suggests that the ionic network plays an important role in the enzymes thermostability, possibly in retaining dimer integrity at high temperatures. Although these data are consistent with previous studies on hyperthermophilic enzymes, without three-dimensional crystal structures of the *Pf*CS ionic network mutants it is impossible to tell whether the site-directed mutation is actually causing the intended disruption to the inter-subunit interactions, or is causing an unexpected disruption to the overall structure of the protein. Kinetic parameters determined for the *Pf*CS mutants suggest that there is no major difference when compared to the *Pf*CS wt, which supports the conclusion that the required mutation has not caused any major disruption to the overall enzyme structure, or to the active site structures at the subunit interface.

It is clear that many mutations made to a thermostable enzyme by site-directed mutagenesis may result in the loss of thermal stability due to unforeseen effects of introducing substitute residues into an enzyme structure and, although determining kinetic parameters can confirm that the catalytic activity has not been greatly affected,

without three-dimensional structures there is no guarantee that any mutation is creating or disrupting the hypothesised interactions.

For these reasons, I wanted to study the role of the ionic network by attempting to increase the thermal stability of a less thermostable enzyme by introducing an ionic network using site-directed mutagenesis.

6.2 Substituting the *Pf*CS equivalent inter-subunit ionic network into *Ta*CS.

Each equivalent residue that contributes to the 5-membered ionic network in *Pf*CS was introduced into the equivalent position, where possible, into *Ta*CS by site-directed mutagenesis. It is understood that the introduction of a charged residue into the structure of an enzyme can cause a disruption until the charged is complemented by the formation of an ionic bond or a network. For this reason each mutation was created singly or in combination with the other amino acid substitutions.

The kinetic data for these *Ta*CS ionic network mutants show comparable values to those of the *Ta*CS wt, yet the thermal inactivation and temperature optima data show a general reduction in thermal stability and thermoactivity for each CS, apart from the *Ta*CS 5sm. From thermal inactivation studies and crystal structure comparison it has been hypothesised that the single D94H mutation introduced into *Ta*CS has created an interaction which, in turn, has stabilised an ionic bond equivalent to the second intersubunit ionic bond found in the more thermostable *Ss*CS.

The *Ta*CS wt and *Pf*CS wt are 42% identical (62% similar), and although this is a fairly high sequence identity, the three-dimensional structure of *Ta*CS wt may not be similar enough to the *Pf*CS wt for the exact positioning of 5 amino acid residues to form an ionic network. However, the *Ss*CS is 59% identical (76% similar) to the *Ta*CS wt sequence, and hence an increase in structural homology may allow the equivalent second ionic bond of *Ss*CS to be successfully substituted into *Ta*CS. Elucidation of the *Ta*CS 5sm three-dimensional structure is required to investigate the nature of the bonds formed from the single mutation to compare with both the *Ta*CS wt and *Ss*CS wt structures.

6.3 Using rational design/site-directed mutagenesis to increase the thermal stability of an enzyme.

Overall, attempts to construct the 5-membered ionic network into *Ta*CS were unsuccessful as both *Ta*CS mutants containing the full ionic networks were very thermolabile compared to the *Ta*CS wt. This may suggest that, although the rational design/site-directed mutagenesis approach is a valuable tool for the study of specific features of thermostability, it may be more beneficial to use directed evolution if the sole purpose is to increase the overall thermostability of an enzyme. It would be interesting to put the *Ta*CS wt through directed evolution with selective pressures for hyperthermostability, to observe what thermally stabilising structural features would evolve, and whether they would involve the formation of an ionic network or employ an alternative stabilising strategy. As discussed before, it seems that hyperthermostable enzymes have employed the use of ionic networks to achieve extreme thermostability, so it would also be interesting to investigate whether an increase in thermostability by

directed evolution would be the result of an increased ionic content. This may be unlikely as the addition of single charged residues into the structure may be disruptive, and hence that mutant may not be selected for during screening for thermal stability. However, if the enzyme already contains an ionic interaction then it may be that a network could form as the charge of each subsequently added charged residue may be 'shared' with neighbouring charged residues, and prevent the disruption caused by a single charged residue.

To summarise, it is clear that both rational design/site-directed mutagenesis and directed evolution are valuable approaches to the study of enzyme thermostability, although it is necessary to select the appropriate method for the required experimental aim. The site-directed mutagenesis method is used to investigate the specific interactions through which nature has achieved protein thermostability. Alternatively, directed evolution can be used to increase the thermostability of an enzyme, not necessarily using the same selective pressures as natural evolution, hence, different strategies of achieving thermostability may be observed.

APPENDIX

Purification Table for *Ta*CS wild-type and the *Ta*CS ionic network mutants

Enzyme	Stage of purification	Volume (ml)	Total activity (U)	Total protein (mg)	Specific activity (U/mg)	Yield (%)	Purification fold
*Ta*CS wt	Cell-free extract	2.0	333.2	15.9	20.9	100.0	-
	Heat treated	2.0	308.7	3.6	85.3	92.6	4.1
	Red A extract	5.0	261.0	2.3	**113.0**	**78.3**	**5.4**
*Ta*CS 3sm	Cell-free extract	2.0	242.4	30.9	7.8	100.0	-
	Heat treated	2.0	243.1	7.7	31.6	100.3	4.0
	Red A extract	5.0	84.0	0.5	**156.2**	**34.7**	**19.9**
*Ta*CS 4sm	Cell-free extract	2.0	171.2	21.4	8.0	100.0	-
	Heat treated	2.0	159.1	3.8	42.3	92.9	5.3
	Red A extract	5.0	95.2	0.8	**118.6**	**55.6**	**14.8**
*Ta*CS 5sm	Cell-free extract	2.0	550.8	35.1	15.7	100.0	-
	Heat treated	2.0	457.6	11.0	41.6	83.1	2.7
	Red A extract	5.0	394.5	3.7	**107.2**	**71.6**	**6.8**
*Ta*CS 3/4dm	Cell-free extract	2.0	499.2	29.8	16.7	100.0	-
	Heat treated	2.0	508.1	9.9	51.1	101.8	3.1
	Red A extract	5.0	223.2	1.0	**216.7**	**44.7**	**12.9**
*Ta*CS 4/5dm	Cell-free extract	2.0	147.2	37.1	4.0	100.0	-
	Heat treated	2.0	150.4	13.4	11.2	102.2	2.8
	Red A extract	5.0	40.9	0.5	**82.6**	**27.8**	**20.8**

Enzyme	Stage of purification	Volume (ml)	Total activity (U)	Total protein (mg)	Specific activity (U/mg)	Yield (%)	Purification fold
*Ta*CS 3/4/5dm	Cell-free extract	2.0	35.1	33.7	1.0	100.0	-
	Heat treated	2.0	41.3	7.7	5.4	117.9	5.2
	Red A extract	5.0	19.6	0.4	**49.1**	**56.0**	**47.1**
*Ta*CS 6sm	Cell-free extract	2.0	347.6	27.8	12.5	100.0	-
	Heat treated	2.0	336.1	12.1	27.9	96.7	2.2
	Red A extract	5.0	88.8	1.0	**90.6**	**25.6**	**7.2**
*Ta*CS 6/3dm	Cell-free extract	2.0	1133	18.7	60.5	100.0	-
	Heat treated	2.0	671.9	8.7	77.2	59.3	1.3
	Red A extract	5.0	61.8	1.3	**49.4**	**5.5**	**0.8**
*Ta*CS 6/5dm	Cell-free extract	2.0	130.8	11.9	11.0	100.0	-
	Heat treated	2.0	185	8.1	23.0	141.5	2.1
	Red A extract	5.0	82.1	0.8	**109.5**	**62.8**	**10**
*Ta*CS 3/5/6tm	Cell-free extract	2.0	71.3	9.3	7.7	100.0	-
	Heat treated	2.0	27.7	3.8	7.3	38.9	1
	Red A extract	5.0	21.8	0.4	**50.1**	**30.5**	**6.5**
Sswt	Cell-free extract	2.0	770.3	30.6	25.2	100.0	-
	Heat treated	2.0	998.8	24.6	40.6	129.7	1.6
	RedA extract	5.0	716.0	5.4	**132.8**	**93**	**5.3**

REFERENCES

Aghajanian, S. & Engel, P.C. (1998). Use of protein engineering to explore subunit interactions in an allosteric enzyme: construction of inter-subunit hybrids in *Clostridium symbiosum* glutamate dehydrogenase. *Protein Eng.* **11**, 569-575.

Alter, G.M., Casazza, J.P., Zhi, W., Nemeth, P., Srere, P.A. & Evans, C.T. (1990). Mutation of Essential Catalytic Residues in Pig Citrate Synthase. *Biochemistry.* **29**, 7557-7563.

Arnold, F.H. & Volkov, A.A. (1999). Directed evolution of biocatalysts. *Curr. Opp. Chem. Biol.* **3**, 54-59.

Arnold, F.H., Wintrode, P.L., Miyazaki, K. & Gershenson, A. (2001). How enzymes adapt: lessons from directed evolution. *Trends Biochem. Sci.* **26**, 100-106.

Arnott, M.A. (1999). Investigation of the Structural Basis of protein Hyperthermostability in Citrate Synthase from the Archaea. PhD thesis, University of Bath.

Arnott, M.A., Michael, R.A., Thompson, C.R., Hough, D.W. & Danson, M.J. (2000). Thermostability and Thermoactivity of Citrate Synthases from the Thermophilic and Hyperthermophilic Archaea, *Thermoplasma acidophilum* and *Pyrococcus furiosus*. *J. Mol. Biol.* **304**, 657-668.

Bell, G.S. (1999). Crystallographic Studies of Central Metabolic Enzymes from Hyperthermophilic Archaea. PhD thesis, University of Bath.

Bogin, O., Peretz, M., Hacham, Y., Korkhin, Y., Frolow, F., Kalb, A.J. & Burstein, Y. (1998). Enhanced thermal stability of *Clostridium beijerinckii* alcohol dehydrogenase after strategic substitution of amino acid residues with prolines from the homologous thermophilic *Thermoanaerobacter brockii* alcohol dehydrogenase. *Protein Sci.* **7**, 1156-1163.

Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analy. Biochem.* **72**, 248-254.

Britton, K.L., Yip, K.S.P., Sedelnikova, S.E., Stillman, T.J., Adams, M.W.W., Ma, K., Maeder, D.L., Robb, F.T., Tolliday, N., Vetriani, C., Rice, D.W. & Baker, P.J. (1999). Structure Determination of the Glutamate Dehydrogenase from the Hyperthermophile *Thermococcus litoralis* and its Comparison with that from *Pyrococcus furiosus*. *J. Mol. Biol.* **293**, 1121-1132.

Brons-Poulsen, J., Peterson, N.E., Horder, M. & Kristiansen, K. (1998). An improved PCR-based method for site directed mutagenesis using megaprimers. *Mol. Cell. Probes.* **12**, 345-348.

Bruins, M.E., Janssen, A.E.M. & Boom, R.M. (2001). Thermozymes and their Applications. *App. Biochem. Biotech.* **90**, 155-186.

Cambillau, C. & Claverie, JM. (2000). Structural and Genomic correlates of Hyperthermostability. *J. Biol. Chem.* **275**, 32383-32386.

Cavagnero, S., Debe, D.A., Zhou, Z.H., Adams, M.W.W. & Chan, S.I. (1998). Kinetic Role of Electrostatic Interactions in the Unfolding of Hyperthermophilic and Mesophilic Rubredoxins. *Biochemistry.* **37**, 3369-3376.

Chen, J., Zhiqiang, L., Sakon, J. & Stites, W.E. (2000). Increasing the Thermostability of Staphylococcal Nuclease: Implications for the Origin of Protein Thermostability. *J. Mol. Biol.* **303**, 125-130.

Chen, R. (2001). Enzyme engineering: rational redesign versus directed evolution. *Trends Biotechnol.* **19**, 13-14.

Connaris, H., West, S.M., Hough, D.W. & Danson, M.J. (1998). Cloning and overexpression in Escherichia coli of the gene encoding citrate synthase from the hyperthermophilic Archaeon *Sulfolobus solfataricus*. *Extremophiles* **2**, 61-66.

Crabb, W.D. & Mitchinson, C. (1997). Enzymes involved in the processing of starch to sugars. *Trends Biotechnol.* **15**, 349-352.

Daniel, R.M., Dines, M. & Petach, H.H. (1996). The denaturation and degradation of stable enzymes at high temperatures. *Biochem. J.* **317**, 1-11.

Daniel, R.M., Danson, M.J. & Eisenthal, R. (2001). The temperature optima of enzymes: a new perspective on an old phenomenon. *Trends Biochem. Sci.* **26**, 223-225.

Danson, M.J., Black, S.C., Woodland, D.L. & Wood, P.A. (1985). Citric acid cycle of the archaebacteria: citrate synthase and succinate thiokinase. *FEBS Lett.* **179**, 120-124.

Danson, M.J., Hough, D.W., Russell, R.J.M., Taylor, G.L. & Pearl, L. (1996). Enzyme thermostability and thermoactivity. *Protein Eng.* **9**, 629-630.

Danson, M.J. & Hough, D.W. (1998). Structure, function and stability of enzymes from the Archaea. *Trends Microbiol.* **6**, 307-314.

Danson, M.J. & Hough, D.W. (2001). Thermostability and Thermoactivity of Extremozymes. *United Nations Educational, Scientific and Cultural Organization (UNESCO)/ The Encyclopedia of Life Support Systems (EOLSS)*, (in press).

Demirjian, D.C., Morís-Varas, F. & Cassidy, C.S. (2001). Enzymes from extremophiles. *Curr. Opp. Chem. Biol.* **5**, 144-151.

Dill, K.A., Alonso, D.O.V. & Hutchinson, K. (1989). Thermal Stabilities of Globular Proteins. *Biochemistry.* **28**, 5439-5449.

Eichler, J. (2001). Biotechnological uses of archaeal extremozymes. *Biotech. Adv.* **19**, 261-278.

Eisenthal, R & Cornish-Bowden, A. (1974). The direct linear plot. A new graphical procedure for estimating enzyme kinetic parameters. *Biochem. J.* **139**, 715-720.

Elcock, A.H. (1998). The Stability of Salt Bridges at High Temperatures: Implications for Hyperthermophilic Proteins. *J. Mol. Biol.* **284**, 489-502.

Erduran, I. & Kocabıyık, S. (1998). Amino acid substitutions in the Subunit Interface Enhancing Thermostability of *Thermoplasma acidophilum* Citrate Synthase. *Biochem. Biophys. Res. Com.* **249**, 566-571.

Evans, C.T., Owens, D., Casazza, J.P. & Srere, P.A. (1989). Studies on site directed mutant pig citrate synthases. *Biochem. Biophys. Res. Comm.* **164**, 1437-1445.

Evans, C.T., Kurz, L.C., Remington, S.J. & Srere, P.A. (1996). Active site mutants of Pig Citrate Synthase: Effects of Mutations on the Enzyme Catalytic and Structural Properties. *Biochemistry.* **35**, 10661-10672.

Eyles, S.J. & Gierasch, L.M. (2000). Multiple Roles of Prolyl residues in Structure and Folding. *J. Mol. Biol.* **301**, 737-747.

Fagain, C.O. (1995). Understanding and increasing protein stability. *Biochim. Biophys. Acta*. **1252**, 1-14.

Fiala, G & Stetter, K.O. (1986). *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C. *Arch. Microbiol.* **145**, 56-61.

Fields, P.A. (2001). Review: Protein function at thermal extremes: balancing stability and flexibility. *Comp. Biochem. and Physiol. A* **129**, 417-431.

Fischer, G. & Bang, H. (1985). The refolding of urea-denatured ribonuclease A is catalysed by peptidyl-prolyl *cis-trans* isomerase. *Biochim. Biophy.s Acta* **198**, 39-42.

Forterre, P. (1997). Archaea: what can we learn from their sequences? *Curr. Opp. Gen. Devel.* **7**, 764-770.

Gerday, C., Aittaleb, M., Bentahir, M., Chessa, JP., Claveria, P., Collins, T., D'Amico, S., Dumont, J., Garsoux, G., Georlette, D., Hoyoux, A., Lonhienne, T., Meuwis, MA. & Feller, G. (2000). Cold-adapted enzymes: from fundamentals to biotechnology. *Trends Biotechnol.* **18**, 103-107.

Gerk, L.P., Leven, O. & Müller-Hill, B. (2000). Strengthening the Dimerisation Interface of Lac Repressor Increases its Thermostability by 40 deg. C. *J. Mol. Biol.* **299**, 805-812.

Gianese, G., Argos, P. & Pascarella, S. (2001). Structural adaptation of enzymes to low temperatures. *Protein Eng.* **14,** 141-148.

Gothel, S.F. & Marahiel, M.A. (1999). Peptidyl-prolyl *cis-trans* isomerases, a superfamily of ubiquitous folding catalysts. *Cell. Mol. Life Sci.* **55**, 423-436.

Hamana, H. & Shinozawa, T. (1999). Effects of C-terminal deletion on the activity and thermostability of orotate phosphoribosyltransferase from *Thermus thermophilus*. *J. Biochem.* **125**, 109-114.

Haney, P.J., Badger, J.H., Buldak, G.L., Reich, C.I., Woese, C.R. & Olsen, G.J. (1999). Thermal adaptation analyzed by comparison of protein sequences from mesophilic and extremely thermophilic *Methanococcus* species. *Proc. Natl. Acad. Sci.* **96,** 3578-3583.

James, K.D., Russell, R.J.M., Parker, L., Daniel, R.M., Hough, D.W. & Danson, M.J. (1994). Citrate synthases from Archaea: development of a bio-specific affinity chromatography purification procedure. *FEMS Microbiol. Letters*, **119**, 181-186.

Karpusas, M., Branchaud, B. & Remington, S.J. (1990). Proposed Mechanism for the Condensation Reaction of Citrate Synthase: 1.9Å Structure of the Ternary Complex with Oxaloacetate and Carboxymethyl Coenzyme A. *Biochemistry.* **29**, 2213-2219.

Karshikoff, K. & Ladenstein. (2001). Ion pairs and the thermotolerance of proteins from hyperthermophiles: a 'traffic rule' for hot rods. *Trends Biochem. Sci.* **9**, 550-556.

Keeling, P.J. & Doolittle, W.F. (1995). Archaea: Narrowing the gap between prokaryotes and eukaryotes. *Proc. Natl. Acad. Sci.* **92**, 5761-5764.

Kumar, S., Ma, B., Tsai, C.J. & Nussinov, R. (2000a). Electrostatic Strengths of Salt Bridges in Thermophilic and Mesophilic Glutamate Dehydrogenase Monomers. *Proteins* **38**, 368-383.

Kumar, S., Tsai, CJ. & Nussinov, R. (2000b). Factors enhancing protein thermostability. *Protein Eng.* **13**, 179-191.

Kurz, L.C., Shah, S., Frieden, C., Nakra, T., Stein, R.E., Drysdale, G.R., Evans, C.T. & Srere, P.A. (1995). Catalytic Stratergy of Citrate Synthase: Subunit Interactions Revealed as a Consequence of a single Amino Acid Change in the Oxaloacetate binding site. *Biochemistry.* **34**, 13278-13288.

Kurz, L.C., Nakra, T., Stein, T., Plungkhen, W., Riley, M., Hsu, F. & Drysdale, G.R. (1998). Effeects of Changes in the Three Catalytic Residues on the Relative stabilities of Some of the Intermediates and Transition States in the Citrate Synthase Reaction. *Biochemistry.* **37**, 9724-9737.

Lebbink., J.H.G., Knapp, S., Van der Oost, J., Rice, D., Ladenstein, R. & De Vos, W.M. (1999). Engineering Activity and Stability of Thermotoga maritima Glutamate Dehydrogenase II: Construction of a 16-Residue Ion-pair Network at the Subunit Interface. *J. Mol. Biol.* **289**, 357-369.

Lehmann, M., Kostrewa, D., Wyss, M., Brugger, R., D'Arcy, A., Pasamontes, L. & van Loon, A.P.G.M. (2000). From DNA sequence to improved functionality: using protein sequence comparisons to rapidly design a thermostable consensus phytase. *Protein Eng.* **13**, 49-57.

Lehmann, M., Pasamontes, L., Lassen, S.F. & Wyss, M. (2000). The consensus consept for thermostability engineering of proteins. *Biochim. Biophys. Acta.* **1543**, 408-415.

Lehmann, M. & Wyss, M. (2001). Engineering proteins for thermostability: the use of sequence alignments versus rational design and directed evolution. *Curr. Opp. Biotech.* **12**, 371-375.

Liu, HL., Doleyers, Y., Coutinho, P.M., Ford, C. & Reilly, P.J. (2000). Replacement and deletion mutations in the catalytic domain and belt region of *Aspergillus awamori* glucoamylase to enhance thermostability. *Protein Eng.* **13**, 655-659.

Lonhienne, T., Gerday, C. & Feller, G. (2000). Psychrophilic enzymes: revisiting the thermodynamic parameters of activation may explain local flexibility. *Biochim. Biophys. Acta* **1543**, 1-10.

Madern, D., Ebel, C. & Zaccai, G. (2000). Halophilic adaptation of enzymes. *Extremophiles* **4**, 91-98.

Man, W.J., Li, Y., O'Connor, D. & Wilton, D.C. (1994). The effect of replacing the conserved active-site residues His-264, Asp-312 and Arg-314 on the binding and catalytic properties of *Escherichia coli* citrate synthase. *Biochem. J.* **300**, 765-770.

Margesin, R. & Schinner, F. (2001). Potential of halotolerant and halophilic microorganisms for biotechnology. *Extremophiles* **5**, 73-83.

Martins, L.O., Huber, R., Huber, H., Stetter, K.O., da Costa, M.S. &Santos,H. (1997). Organic Solutes in Hyperthermophilic Archaea. *Appl. Environ. Microbiol.* **63**, 869-902.

Maruyama, T. & Furutani, M. (2000). Archaeal Peptidyl prolyl *cis-trans* isomerases (PPIases). *Frontiers in Bioscience* **5**, 821-836.

Merz, A., Knöchel, T., Jansonius, J.N. & Kirschner K. (1999). The Hyperthermostable Indoleglycerol Phosphate Synthase from *Thermotoga maritima* is Destabilized by Mutational Disruption of Two Solvent-exposed Salt Bridges. *J. Mol. Biol.* **288**, 753-763.

Mevarech, M., Frolow, F. & Gloss, L.M. (2000). Halophilic enzymes: proteins with a grain of salt. *Biophys. Chem.* **86**, 155-164.

Miyazaki, K., Wintrode, P.L., Grayling, R.A., Rubingh, D.N. & Arnold, F.H. (2000). Directed Evolution Study of Temperature Adaptation in a Psychrophilic Enzyme. *J. Mol. Biol.* **297**, 1015-1026.

Moore, J.C., Jin, H.M., Kuchner, O. & Arnold, F.H. (1997). Strategies for the *in vitro* Evolution of Protein Function: Enzyme Evolution by Random recombination of Improved Sequences. *J. Mol. Biol.* **272**, 336-347.

Muir, J.M., Russell, R.J.M., Hough, D.W. & Danson, M.J. (1995). Citrate synthase from the hyperthermopilic Archaeon, *Pyrococcus furiosus*. *Protein Eng.* **8**, 583-592.

Mulholland, A.J. & Richards, W.G. (1998a). A model of the condensation step in the citrate synthase reaction. *J.Mol. Structure (Theochem)* **427**, 175-184.

Mulholland, A.J. & Richards, W.G. (1998b). Modeling Enzyme Reaction Intermediates and Transition States: Citrate Synthase. *J. Phys. Chem. B* **102**, 6635-6646.

Nèmeth, A., Svingor, Á., Pócsik, M., Dobó, J., Magyar, C, Szilágyi, A., Gál, P. & Závodszky, P. (2000). Mirror image mutations reveal the significance of an intersubunit ion cluster in the stability of 3-isopropylmalate dehydrogenase. *FEBS Lett.* **468**, 48-52.

Pal. D. & Chakrabati, P. (1999). *Cis* Peptide Bonds in Proteins: Residues Involved, their Conformations, Interactions and Locations. *J. Mol. Biol.* **294**, 271-288.

Pappenberger, G., Schurig, H. & Jaenicke, R. (1997). Disruption of an Ionic Network Leads to Accelerated Thermal Denaturation Of D-Glyceraldehyde-3-Phospate Dehydrogenase from the Hyperthermophilic Bacterium *Thermotoga maritima*. *J. Mol. Biol.* **274**, 676-683.

Petrounia, I.P. & Arnold, F.H. (2000). Designed evolution of enzymatic properties. *Curr. Opp. Biotech.* **11**, 325-330.

Rahman, R.N.Z.A., Fujiwara, S., Nakamura, H., Takagi, M. & Imanaka, T. (1998). Ion Pairs Involved in Maintaining a Thermostable Structure of Glutamate Dehydrogenase from a Hyperthermophilic Archaeon. *Biochem. Biophys. Res. Comm.* **248**, 920-926.

Rees, D.C. & Robertson, A.D. (2001). Some thermodynamic implications for the thermostability of proteins. *Protein Sci.* **10**, 1187-1194.

Remington, S., Wiegand, G. & Huber, R. (1982). Crystallographic Refinement and Atomic Models of Two Different Forms of Citrate Synthase at 2.7 and 1.7Å resolution. *J. Mol. Biol.* **158**, 111-152.

Remington, S.J. (1992). Structure and Mechanism of Citrate Synthase. *Curr. Top. Cell. Regul.* **33**, 209-229.

Robb, F.T., Maeder, D.L., Brown, J.R., DiRuggiero, J., Stump, M.D., Yeh, R.K., Weiss, R.B. & Dunn, D.M. (2001). Genomic sequence of hyperthermophile, *Pyrococcus furiosus*; implications for physiology and enzymology. *Methods Enzymol.* **330**, 134-157.

Ruepp, A., Graml, W., Santos-Martinez, ML., Koretke, K.K., Volker, C., Mewes, H.W., Frishman, D., Stocker, S., Lupas, A.N. & Baumeister, W. (2000). The genome sequence of the thermoacidophilic scavenger *Thermoplasma acidophilum*. *Nature* **407**, 508-513.

Russell, R.J.M., Hough, D.W., Danson, M.J. & Taylor, G.L. (1994). The crystal structure of citrate synthase from the thermophilic Archaeon, *Thermoplasma acidophilum*. *Structure* **2**, 1157-1167.

Russell, R.J.M., Ferguson, M.C., Hough, D.W., Danson, M.J. & Taylor, G.L. (1997). The Crystal Structure of Citrate Synthase from the Hyperthermophilic Archaeon *Pyrococcus furiosus* at 1.9Å Resolution. *Biochemistry.* **36**, 9983-9994.

Russell, R.J.M., Gerike, U., Danson, M.J., Hough, D.W. & Taylor, G,L. (1998). Structural adaptations of the cold-active citrate synthase from an Antartic bacterium. *Structure* **6**, 351-361.

Schiene, C., Reimer, U., Schutkowski, M. & Fischer, G. (1998). Mapping the stereospecificity of peptidyl prolyl *cis/trans* isomerases. *FEBS Lett.* **432**, 202-206.

She, Q., Singh, R.K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M.J., Chan-Weiher, C.C.Y., Clausen, I.G., Curtis, B.A., De Moors, A., Erauso, G., Fletcher, C., Gordon, P.M.K., Jong, I.H., Jeffries, A.C., Kozera, C.J., Medina, N., Peng, X., Thi-Ngoc, H.P., Duguet, M., Gaasterland, T., Garrett, R.A., ragan, M.A., Sensen, C.W. & Van der Oost, J. (2001). The complete genome of the crenarchaeon *Sulfolobus solfataricus*. *Proc. Natl. Acad. Sci.* **98**, 7835-7840.

Smith, L.D., Stevenson, K.J., Hough, D.W. & Danson, M.J. (1987). Citrate synthase from the thermophilic archaebacteria *Thermoplasma acidophilum* and *Sulfolobus solfataricus*. *FEBS Lett.* **225**, 277-281.

Srere, P.A., Brazil, H. & Gonen, L. (1963). The citrate condensing enzyme of pigeon breast muscle and moth flight muscle. *Acta Chem, Scand.* **17**, S129-S134.

Sriprapundh, D., Vieille, C. & Zeikus, J.G. (2000). Molecular determinants of xylose isomerase thermal stability and activity: analysis of thermozymes by site-directed mutagenesis. *Protein Eng.* **13**, 259-265.

Sterner, R. & Liebl, W. (2001). Thermophilic adaptation of Proteins. *Crit. Rev. Biochem. Mol. Biol.* **36**, 39-106.

Stewart, D.E., Sarkar, A. & Wampler, J.E. (1990). Occurence and role of *cis* peptide bonds in protein structures. *J. Mol. Biol.* **214**, 253-260.

Sun, M.M.C., Tolliday, N., Vetriani, C., Robb, F.T. & Clark, D.S. (1999). Pressure-induced thermostabilization of glutamate dehydrogenase from hyperthermophile *Pyrococcus furiosus*. *Protein Sci.* **8**, 1056-1063.

Sutherland, K.J., Henneke, C.M., Towner, P., Hough, D.W. & Danson, M.J. (1990). Citrate synthase from the thermophilic archaebacterium *Thermoplasma acidophilum*: Cloning and sequencing of the gene. *Eur. J. Biochem.* **194**, 839-844.

Sutherland, K.J., Danson, M.J., Hough, D.W. & Towner, P. (1991). Expression and purification of plasmid-encoded *Thermoplasma acidophilum* citrate synthase from *Escherichia coli*. *FEBS Lett.* **282**, 132-4.

Svingor, A., Kardos, J., Hajdú, I., Nèmeth, A. & Závodszky, P. (2001). A Better Enzyme to Cope with Cold: Comparative flexibility studies on psychrotrophic, mesophilic and thermophilic IPMDHs. *J. Biol. Chem.* **276**, 28121-28125.

Szilágyi, A. & Závodszky, P. (2000). Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. *Structure* **8**, 493-504.

Thompson, M.J. & Eisenberg, D. (1999). Transproteomic Evidence of a Loop-Deletion Mechanism for Enhancing Protein Thermostability. *J. Mol. Biol.* **290**, 595-604.

Ursby, T., Adinolfi, B.S., Al-Karadaghi, S., De Vendittis, E. & Bocchini, V. (1999). Iron Superoxide Dismutase from the Archaeon *Sulfolobus solfatricus*: Analysis of Structure and Thermostability. *J. Mol. Biol.* **286**, 189-205.

Van den Burg, B., Vriend, G., Veltman, O.R., Venema, G. & Eijsink, V.G.H. (1998). Engineering an enzyme to resist boiling. *Proc. Natl. Acad. Sci.* **95**, 2056-2060.

Vetriani, C., Maeder, D.L., Tolliday, N., Yip, K.S.P., Stillman, T.J., Britton, K.L., Rice, D.W., Klump, H.H. Robb, F.T. (1998). Protein thermostability above 100°C: A key role for ionic interactions. *Proc. Natl. Acad. Sci.* **95**, 12300-12305.

Vieille, C. & Zeikus, G.J. (1996). Thermozymes: identifying molecular determinants of protein structural and functional stability. *Trends Biotechnol.* **14**, 183-190.

Vieille, C. & Zeikus, G.J. (2001). Hyperthermophilic Enzymes: Sources, Uses, and Molecular Mechanisms for Thermostability. *Microbiol. Mol. Biol. Rev.* **65**, 1-43.

Vogt, G., Woell, S. & Argos, P. (1997). Protein Thermal Stability, Hydrogen Bonds, and Ion Pairs. *J. Mol. Biol.* **269**, 631-643.

Von Ahsen, O., Lim, J.H., Caspers, P., Martin, F., Schonfeld, H-J., Rassow, J & Pfanner, N. (2000). Cyclophilin-promoted folding of mouse dihydrofolate reductase does not include the slow conversion of the late-folding intermediate to the active enzyme. *J. Mol. Biol.* **294**, 809-818.

Wallace, L.A., Burke, J. & Dirr, H.W. (2000). Domain-domain interface packing at conserved Trp-20 in class α glutathione transferase impacts on protein stability. *Biochim.Biophys. Acta.* **1478**, 325-332.

Watanabe, K. & Suzuki, Y. (1998). Protein thermostabilization by proline substitutions. *J.Mol. Cat. B* **4**, 167-180.

Wintrode, P.L., Miyazaki, K. & Arnold, F.H. (2001). Patterns of adaption in a laboratory evolved thermophilic enzyme. *Biochim. Biophys. Acta.* **1549**, 1-8.

Woese, C.R., Kandler, O. & Wheelis, M.L. (1990). Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci.* **87**, 4576-4579.

Yamataga, Y., Ogasahara, K., Hioki, Y., Lee, S.J., Nakagawa, A., Nakamura, H., Ishida, M., Kuramitsu, S. & Yutani, K. (2001). Entropic Stabilization of the Tryptophan Synthase α-Subunit from a Hyperthermophile, *Pyrococcus furiosus*. *J.Biol. Chem.* **276**, 11062-11071.

Yip, K.S.P., Stillman, T.J., Britton, K.L., Artymiuk, P.J., Baker, P.J., Sedelnikova, S.E., Engel, P.C., Pasquo, A., Chiaraluce, R., Consalvi, V., Scandurra, R. & Rice, D.W. (1995). The structure of *Pyrococcus furiosus* glutamate dehydrogenase reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures. *Structure* **3**, 1147-1158.

Yip, K.S.P., Britton, L., Stillman, T.J., Lebbink, J., De Vos, W.M., Robb, F.T., Vetriani, C., Maeder, D. & Rice, D.W. (1998). Insights into the molecular basis of thermal stability from the analysis of ion-pair networks in the glutamate dehydrogenase family. *Eur. J. Biochem.* **255**, 336-346.

Závodszky, P., Kardos, J., Svingor, Á. & Petsko, G. (1998). Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc. Natl. Acad. Sci.* **95**, 7406-7411.

Zhi, W., Srere, P. & Evans, C.T. (1991). Conformational Stability of Pig Citrate Synthase and Some Active-Site Mutants. *Biochemistry.* **30**, 9281-9286.